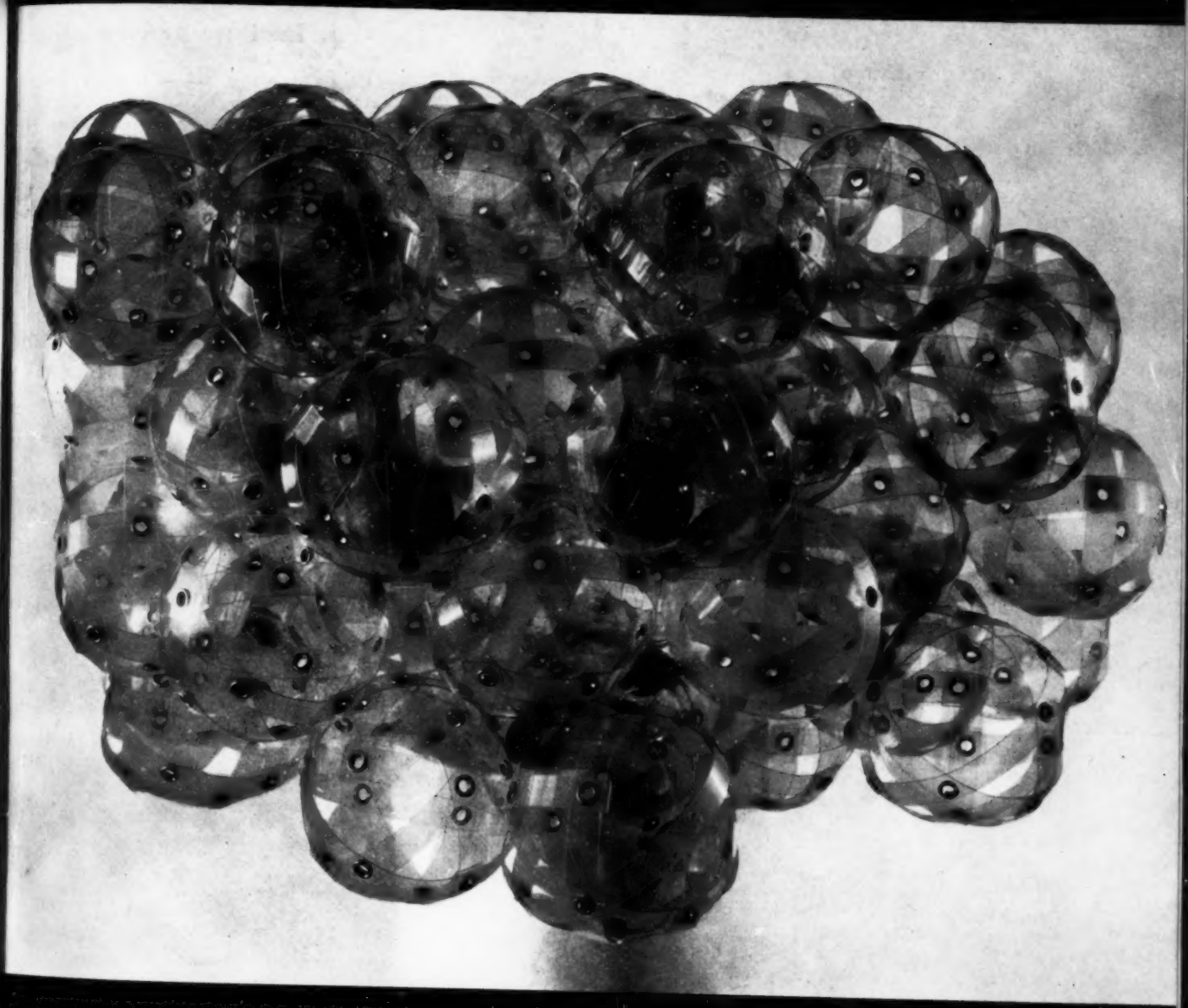
# JOURNAL OF APPLIED PHYSICS



JOURNAL OF APPLIED

# PHYSICS

(Formerly PHYSICS)

APRIL, 1938 Vol. 9, No. 4

## *In This Issue*

*Page*

**Polarized Light Enters the World of Everyday Life** MARTIN GRABAU 215

**Frederic Eugene Ives** EDWARD EPSTEAN AND J. A. TENNANT 226

**X-Rays in Agriculture** LOUIS R. MAXWELL AND STERLING B. HENDRICKS 237

**Order and Disorder in Alloys** H. A. BETHE 244

**Reaction Rate Theory of Viscosity and Some of Its Applications** RAYMOND H. EWELL 252

*Contributed Original Research*

**Reflection of Plane Wave Pulses from Plane Parallel Plates** MORRIS MUSKAT 275

**Effect of Pressure Upon the Elastic Parameters of Isotropic Solids, According to Murnaghan's Theory of Finite Strain** FRANCIS BIRCH 279

*In Every Issue*

**Here and There** 270
**Innovations in Instruments** 271
**Some of the New Books** 272
**Among Our Authors** 274
**Current Literature of Physics** viii
**Advertising Index** xxii

**BOARD OF EDITORS**

ELMER HUTCHISSON, *Editor*

*Associate Editors*

R. M. BOZORTH
D. W. BRONK
W. P. DAVEY
O. S. DUFFENDACK
SAUL DUSHMAN
A. STUART HUNTER
C. H. KUNSMAN
J. T. LITTLETON
L. L. NETTLETON
W. B. NOTTINGHAM
A. R. OLPIN
JOHN T. TATE

**Publication of the American Institute of Physics**

**ADMINISTRATIVE STAFF**

HENRY A. BARTON, *Director*

MADELINE M. MITCHELL, *Publications Manager*

WILLIAM C. COPP, *Advertising Manager*

**GOVERNING BOARD**

JOHN T. TATE, *Chairman*
KARL T. COMPTON
E. C. CRITTENDEN
WHEELER P. DAVEY
HOMER L. DODGE
HARVEY FLETCHER
ARTHUR C. HARDY
VERN O. KNUDSEN
MELVIN MOONEY
FREDERIC PALMER, JR.
R. L. PEEK, JR.
GEORGE B. PEGRAM
W. B. RAYTON
F. K. RICHTMYER
WALLACE WATERFALL

THE JOURNAL OF APPLIED PHYSICS, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as PHYSICS. It aims to be of service not only in physical laboratories, but also in the laboratories of industrial, chemical, geological, geophysical, meteorological, radio, and similar concerns. THE JOURNAL OF APPLIED PHYSICS publishes editorials, news of physicists and reviews, as well as technical papers of applied physics.

SUBSCRIPTION PRICE

| | U. S., POSS. AND CANADA | FOREIGN |
|---|---|---|
| To members of scientific societies | $5.00 | $5.70 |
| To all others | 7.00 | 7.70 |

**Manuscripts** should be submitted to Elmer Hutchisson, Editor, University of Pittsburgh, Pittsburgh, Pennsylvania.

**Changes of address,** proofs, and complaints of failure to receive the JOURNAL OF APPLIED PHYSICS should be addressed to the Publications Manager. New copies can be sent free only if notice of nondelivery is received within three months of date of issue.

**Subscriptions** and orders for back numbers should be addressed to Prince and Lemon Streets, Lancaster, Pa. or to the American Institute of Physics, 175 Fifth Avenue, New York, N. Y.

**Advertising rates** supplied on request. Orders, advertising copy and cuts should be sent to the American Institute of Physics.

Entered as second class matter at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1897. Accepted for mailing at the special rate of postage provided for in the Act of February 28, 1925, authorized May 2, 1932.

Copyright 1938, by the American Institute of Physics.

Journal of Applied Physics

Volume 9 April, 1938 Number 4

# Polarized Light Enters the World of Everyday Life

By Martin Grabau

*The Land-Wheelwright Laboratories, Inc., Boston, Massachusetts*

AN extension of the field of applied optics is now under way whose general importance may ultimately be comparable to the developments which the invention of the vacuum tube has wrought in applied electricity. This new departure revolves about the recent development of an inexpensive, optical polarizing material which releases the phenomenon of polarization of light for practical applications on a large scale.

Apart from colorimetry, applied optics has until now concerned itself with little more than the design of lenses, optical instruments, and illumination systems. Except in such restricted applications as photoelasticity, polarimetry, etc., the essentially wave-like character of light has not been generally useful.

The development of practical applications of wave optics and, in particular, of polarized light was impeded by a formidable obstacle. Polarization was well known and understood, but no cheap and convenient means of polarizing light were available. The Nicol prism and similar contrivances are very nearly perfect polarizers, but they are expensive and limited in size. Polarization of light by reflection from glass at Brewster's angle or by transmission through an inclined pile of glass plates is cheap and simple enough; but both methods are cumbersome and otherwise unsatisfactory. Brewsterian reflection is wasteful of light, and polarization by transmission yields too imperfect a degree of polarization.

A shadow in polarized light.

During the past few years, three new polarizing materials have become commercially available, *viz.*: the Marks Plates,[1] the Zeiss Bernotar filters,[2] and Polaroid.[3] Each of these materials in one way or another utilizes the so-called dichroic properties of certain crystals. The phenomenon of dichroism has been known since 1815 when

Biot observed the polarizing properties of tourmaline. Until now, however, it has remained almost exclusively the possession of the mineralogists. In the literature of physics it is mentioned only casually in connection with the polarization of light by tourmaline.

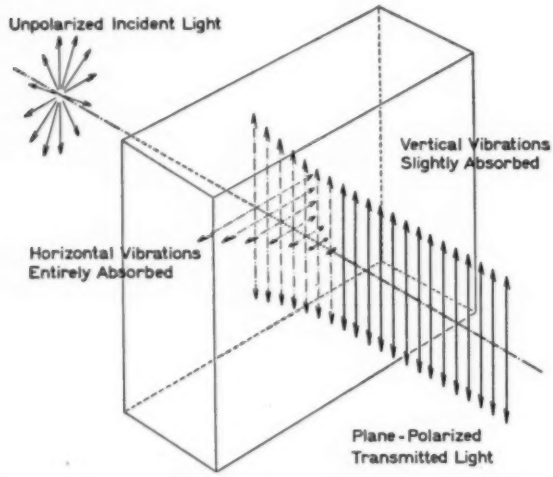



FIG. 1. Diagram illustrating the production of polarized light by a dichroic crystal.

## DICHROISM

It is a well-known fact that a beam of light entering an anisotropic crystal is generally divided into two components polarized at right angles to each other. The wave velocities of these two components are different, so that each direction of propagation in the crystal is, in general, identified with two different indices of refraction. These velocities vary with the orientation of the rays in the crystal. The most general type of variation is found in so-called biaxial crystals. Sometimes, however, it exhibits a certain symmetry, and then the crystal is said to be uniaxial.

Since there are forms of matter in which the index of refraction depends upon the polarization of the light, as in any anisotropic crystal, it is not surprising that there are crystals in which the *absorption* of light is similarly governed by the polarization. The absorption of light by such crystals may be described by an analysis very similar to the treatment of refraction. When the absorption exhibits a degree of generality analogous to the refraction of light in a biaxial crystal, the crystal is said to be *pleochroic;* and when pleochroism exhibits a certain symmetry, analogous to the birefringence of a uniaxial crystal, it degenerates into *dichroism.*

The behavior of a dichroic crystal is illustrated in Fig. 1. A beam of unpolarized light originates in the upper left-hand corner of the diagram. As it enters the crystal, it is divided into two components polarized (as drawn) vertically and horizontally. In the diagram, the crystal transmits the vertical vibrations with very little absorption, while the horizontal vibrations are strongly absorbed. If the attenuation of the latter is sufficiently great, the emergent light is substantially free of horizontal vibrations and is therefore plane polarized. With the formation of a polarizing sheet in mind, the more freely transmitted component may be regarded as the *desired* component, while the other is the *undesired* component. Since the crystal in Fig. 1 is assumed dichroic, it can be rotated in the diagram about either the vertical or horizontal axis (as the case may be) without altering the specific absorptions of the two components.*

Tourmaline is probably the best known dichroic substance. Unfortunately, its absorption coefficients are such that a section sufficiently thick to extinguish the undesired component also reduces the intensity of the desired component unduly. Furthermore, the spectral distribution of transmission in the desired component is very selective, so that the transmitted light is strongly colored. Besides, it is surprisingly difficult to find two pieces of tourmaline with identical optical properties, except when they are cut from the same parent crystal.

The new polarizing materials all utilize the dichroic properties of sulphate of iodo-quinine or related substances. The extraordinary properties of sulphate of iodo-quinine were discovered in 1852 by Herapath.[4] This substance, which is now known as herapathite, possesses the exceedingly useful property by virtue of which a relatively thin section absorbs all visible light in the undesired component almost completely. At the same time it transmits the desired component freely, without undue wave-length selectivity. The crystallography of herapathite has been described by West.[5] Herapath well knew the great value of his discovery, but he found his precious crystals so

* There are certain exceptions to this statement that do not apply to the polarizing crystals in the new polarizers.

fragile and unstable that he despaired of developing the art of growing them into polarizing sheets of large area.

## The New Polarizers

In forming a polarizer of dichroic material, one can either try to perfect Herapath's unsuccessful technique of growing single crystals in sheet form, or else go to the other extreme by assembling a great number of very small crystals, uniformly oriented, in or on a transparent support.

Bernauer's Bernotar filters (Zeiss) are an example of the former approach. With exceedingly great skill and patience, he developed the art of growing single crystal sheets of herapathite only a few thousandths of an inch thick and of mounting them between plates of glass. Despite the logical elegance of Bernauer's process, it hardly lends itself to the manufacture of polarizing areas on such a scale as is contemplated in many of the prospective applications. Its limitations have been indicated by Roesch.[6]

The Marks Polarizing Plates consist of many flat dichroic crystals deposited on the surface of glass, their polarizing axes being substantially parallel. The process of manufacture disclosed in the patent[1] is slow. The patent cites one instance in which forty-five days are required to produce a polarizer 69 mm wide. A similar method of depositing oriented crystals of herapathite on glass to form a polarizing medium has been described by Zimmern and Coutin.[7]

In Polaroid, the small, needle-shaped dichroic crystals are dispersed in great numbers, with substantial uniformity of orientation, in a cellulosic film. The average size of the crystals, their numerical density, and the thickness of the film are easily controlled in a completely mechanized process that operates quickly and continuously. In order to protect the film from dirt and damage, it is then mounted on or between sheets of glass or other transparent materials. There is hardly any practical limit to the width in which the continuous film can be produced. Sheets thirty inches wide have long been on public display. At the present time, Polaroid is the only polarizing material that is commercially available in large sizes and quantities. For a further discussion of Polaroid, the reader is referred to recent articles by Pollard[8] and by Freundlich.[9]

## The Properties of Polaroid

The individual crystals in the Polaroid film are exceedingly small; their girth diameter is considerably less than a wave-length of visible light. Their numerical density is so great that any given ray must traverse a large number of crystals in succession; and the ensemble exhibits no clear spaces whatever. The appearance of the film is completely homogeneous, except under special conditions of polarized illumination at magnifications greater than 1100. The film does not deteriorate with time, and it readily withstands exposure to heat up to 250°F.

The optical properties of Polaroid vary with the numerical density of the crystals and the thickness of the film. Films have been produced varrying in thickness from 0.0001 to 0.01 inch. As the thickness is increased, other factors being constant, the transmission coefficient decreases and the degree of polarization* of the transmitted light increases. Fig. 2 shows the spectral distribu-

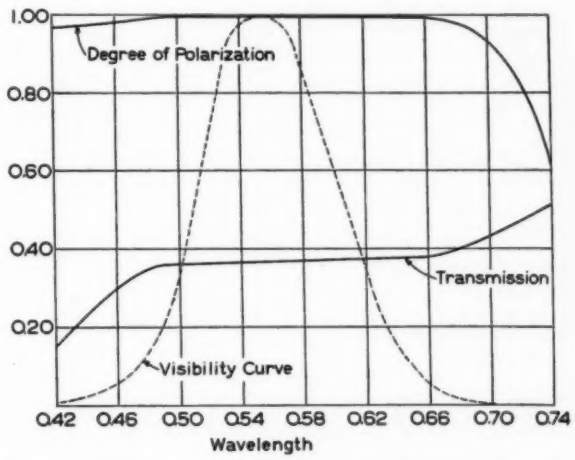



Fig. 2. The spectral distribution of transmission and polarizing power of commercial Polaroid (July 1937).

tion, for visible light, of the relative intensity and the degree of polarization of light transmitted by commercial Polaroid (July 1937) made with film 0.0026 inch thick.[10] The dotted line is the visibility curve of the human eye. It is seen at once that the common usefulness of Polaroid is limited to visible light. At the blue threshold of the visible spectrum, the transmission is well on

* The degree of polarization is defined as the ratio of the intensity of the plane polarized part to the intensity of the transmitted light as a whole.

its way toward ultimate extinction in the near ultraviolet; and at the red end of the spectrum, the degree of polarization is rapidly deteriorating.

A perfect polarizer would transmit fifty percent of the incident light, but the inevitable reflection losses at the front and back surfaces would im-

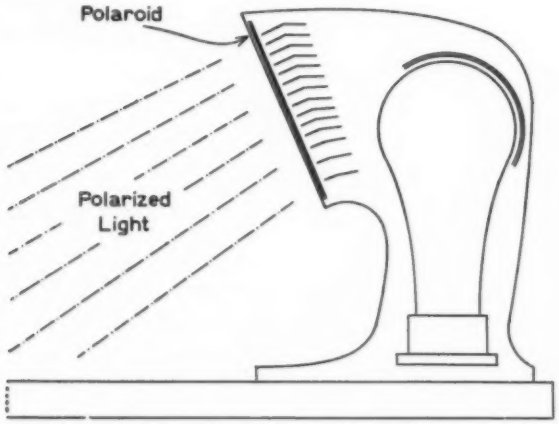



FIG. 3. Cross section of one type of Polaroid desk lamp.

mediately reduce this to forty-six percent. Since the effective average transmission of Polaroid is thirty-seven percent, its efficiency as a transmitter of polarized light is roughly eighty percent. The transmission of a Glan-Thompson prism is slightly less than forty percent,[11] so that its efficiency is approximately eighty-six percent.

It is important to note that both the transmission and polarizing power of commercial Polaroid are nearly constant throughout the central part of the visible spectrum (from 4800 to 6800A) within which the sensitivity of the eye is largely concentrated. In this region, the degree of polarization is greater than 0.996.

The use of thinner films of the same concentration yields higher coefficients of transmission, albeit at the expense of polarizing power. Transmissions greater than forty percent, without too serious losses in percentage polarization (for many practical applications) are easily attainable.

## CLASSIFIED LIST OF SOME APPLICATIONS OF PLANE POLARIZED LIGHT

### *I. Utilization of natural polarization*

Every application of polarized light requires the action of two polarizing processes, one to polarize the light initially and the other to recognize the fact that it is polarized. Nature, however, provides two common sources of polarized light—reflection and scattering—which give rise to some important applications involving only a single sheet of polarizing material.

Whenever light is reflected from a nonmetallic surface (except at normal or grazing incidence), the process of reflection generally involves a differential loss of intensity between the electric vibrations in and normal to the plane of incidence. This is tantamount to saying that the process of reflection alters the state of polarization of the light. If the incident light is unpolarized, the reflected light is at least partially polarized. In the case of vitreous reflection at Brewster's angle,* the reflected light is completely plane polarized.

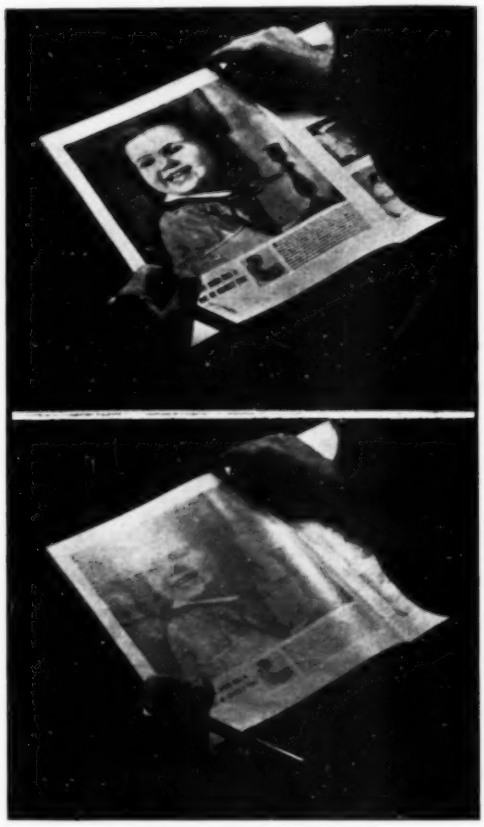


FIG. 4. Magazine page seen under controlled, polarized illumination and the same page as seen when the illumination is "depolarized" by interposing a suitably oriented sheet of Cellophane.

* The tangent of Brewster's angle is equal to the index of refraction of the reflecting medium.

Certain types of gloss meters used in the paper industry are merely devices for measuring the degree of polarization of reflected light.

The analysis of the reflection of light from a mat surface, such as paper, is not as simple as one might expect.[12] For ordinary practical purposes, however, it may be regarded as a combination of quasi-vitreous and of "diffuse" reflection, the latter being entirely unpolarized. Over a surprisingly large range of angles, the so-called glare-light reflected from a mat surface exhibits considerable polarization. For instance, in light reflected from glossy black papers, the degree of polarization may rise to 0.75. In all cases, the predominant electric vibration in the reflected light is normal to the plane of incidence, or parallel to the reflecting surface. The first reference to this important phenomenon occurs in a relatively obscure communication by Malus.[13]

In the case of metallic reflection, the alteration of the state of polarization also involves differential changes of phase, and these are usually just as important as the losses of intensity of the components in and normal to the plane of incidence. Cases of this sort cannot be dealt with as easily as those involving mat reflections. Surface films on the metal usually complicate matters still more.

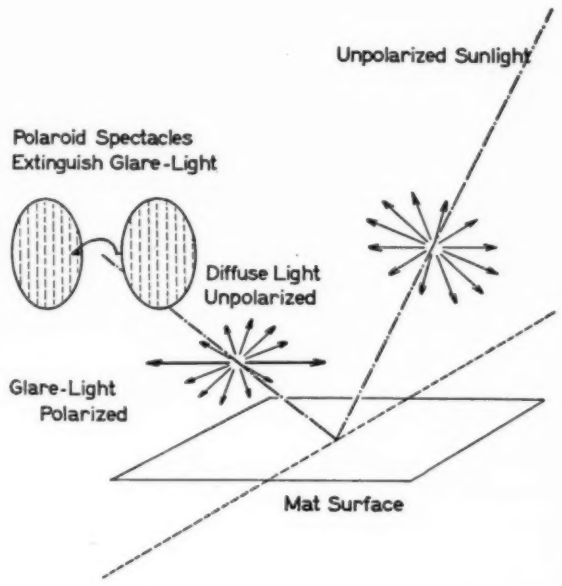



FIG. 5. Diagram illustrating the principle of polarizing sunglasses.

## DESK LAMPS

Under ordinary illumination, the glare from a sheet of paper lying horizontally on a table is largely composed of horizontal vibrations. In order to eliminate the glare, it is primarily necessary to suppress the horizontal vibrations of light at the source.

FIG. 6. Scene in Kenmore Square, Boston, seen without and with Polaroid sunglasses.

In practice, if a desk lamp is so designed that all the light strikes the horizontal paper surface at angles within an easily controllable range, then, by covering the opening of the lamp with a sheet of Polaroid so oriented as to eliminate the horizontal (electric) vibrations that are favored in the reflection process, the glare from the paper (including the inclined facets in its texture) is substantially eliminated. Fig. 3 shows a schematic cross section of one of several types of lamps that satisfy this condition. In this particular design, the inclinations of the rays are controlled by suitably disposed louvers.

## SUNGLASSES

Almost all the glare seen on bright days originates in reflections of sunlight from such horizontal mat surfaces as sidewalks, street pavements, etc. This glare-light is predominantly polarized in the horizontal plane. Sunlight reflected from the sea, ice, and glazed snow is similarly affected. A pair of polarizing spectacles, with vertical polarizing axes, quenches the glare very considerably and creates a freedom from ocular fatigue that is difficult to describe. The principle of this application is illustrated schematically in Fig. 5.

### Filters for Visual Instruments

The use of spyglasses and sextants at sea has always been attended by the disturbing glare of sunlight reflected from the water. Hulburt[14] has observed the high percentage polarization of this glare-light and shown how it can be quenched with suitably oriented polarizing filters. In the case of the sextant, the battery of colored filters now used to absorb direct light from the sun may be replaced by two superposed polarizers, one being rotatable. In order to absorb the far red and infra-red light, the ensemble may be reenforced by a single sheet of suitably absorbent glass. The users of ordinary binoculars will also find polarizing filters most effective on many occasions.

### Polarizing Filters for Photography

In scenes to be photographed, there is often a certain surface reflection that creates a disturbing highlight; and over quite a large range of angles, this light will have undergone considerable polarization. A polarizing filter over the camera lens brings such a highlight under control, more or less independently of the rest of the picture, and makes it possible to bring out the detail of the picture over a wide range of intensities, as it were to circumvent the relatively short latitude of the photographic emulsion.

Another important photographic application is that of a sky filter. The process of scattering of light that generates the blue color of the sky also produces various degrees of polarization. By so orienting a polarizing filter as to reduce the intensity of sky light admitted to the camera, the cloud-to-sky contrast is intensified. This application of a polarizing filter is especially important in color photography where the use of ordinary color filters is out of the question.

The use of polarizing filters in photography has been described at length by Tuttle and McFarlane[15] and by Alfieri.[16]

### Museum and Display Illumination

An extension of the technique developed for the design of Polaroid desk lamps can also be applied to the problems of museum and display illumination. For instance, an oil painting illuminated with suitably disposed polarized light appears to far better advantage than when seen under ordinary light. The surface reflections from varnish and pigments are subordinated to the advantage of genuinely colored light scattered within and beneath the particles of pigment. This enhances the saturations of the colors and enlivens the contrasts considerably.

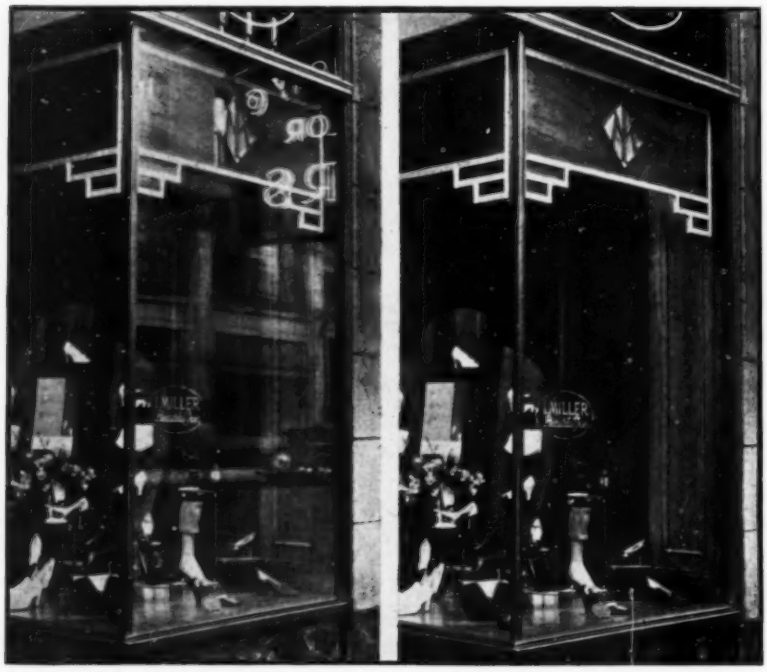


Fig. 7. Photographs of a shop window taken without and with a polarizing filter over the camera lens.

## *II. Use of extinction of light by "crossed" polarizers*

### Elimination of Automobile Headlight Glare

The sketch in Fig. 8 shows one of the ways in which polarization of light can be applied to eliminate automobile headlight glare.[17] Every vehicle may have its headlamps and windshield equipped with Polaroid screens. In the arrangement shown in Fig. 8 the axes of the polarizing

screens on any one vehicle will be uniformly oriented at 45° from the vertical. The driver can then see ahead in the light of his own headlamps, but his windshield screen will quench the direct light from the headlamps of a car coming in the opposite direction.

The photographs in Fig. 9 below will serve to illustrate the effect of this procedure in practice. The headlights of the automobile were covered

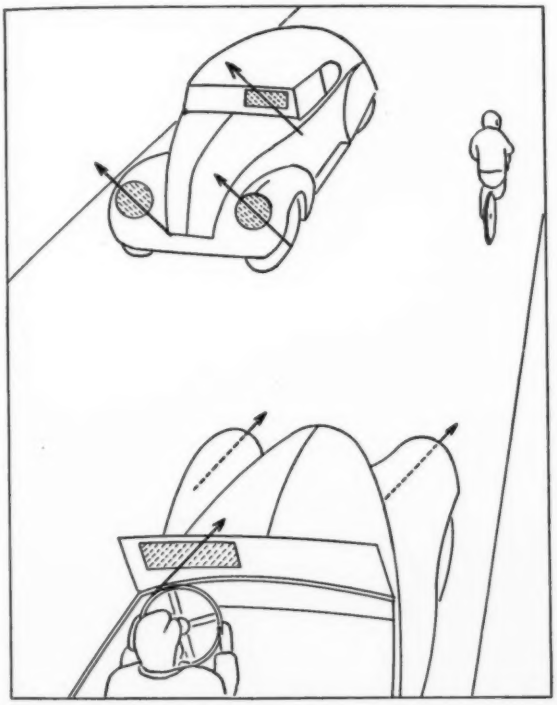


FIG. 8. Sketch illustrating a method of eliminating automobile headlight glare with polarizing screens.

with identically oriented sheets of Polaroid and pointed at the camera. At the same time, the automobile was illuminated by the headlights of another motor car standing to the right of the camera. For the upper picture, taken without a polarizing filter over the camera lens, the time of exposure was necessarily very short. For the lower picture, the camera lens was covered with a Polaroid filter oriented to quench the direct headlight; and the time of exposure was about five hundred times as long as the other. In both cases, a minus-blue filter was used to compensate for the disparity between the blue sensitivity of the eye and of the photographic film. Much of the apparent brightness of the headlights in the lower picture is due to light reflected from the outer surfaces.

However, this photograph can hardly convey the astonishing impression created by being able to read the number plate between the headlights and to see the occupants of an automobile approaching head-on. At the same time, the illumination of the road ahead remains as effective as it was before the approaching car came into view.

## PHOTOMETRICAL USES IN OPTICAL INSTRUMENTS

When one of two superposed sheets of Polaroid is rotated, the relative intensity of visible light transmitted by the ensemble follows the familiar cosine-squared law very closely.[10] The error is relatively very small except within a few degrees of the extinction setting where the effect of the slight defect of polarization becomes relatively large. Such a combination yields a range of relative intensities of at least two hundred to one; and this range can be extended enormously by

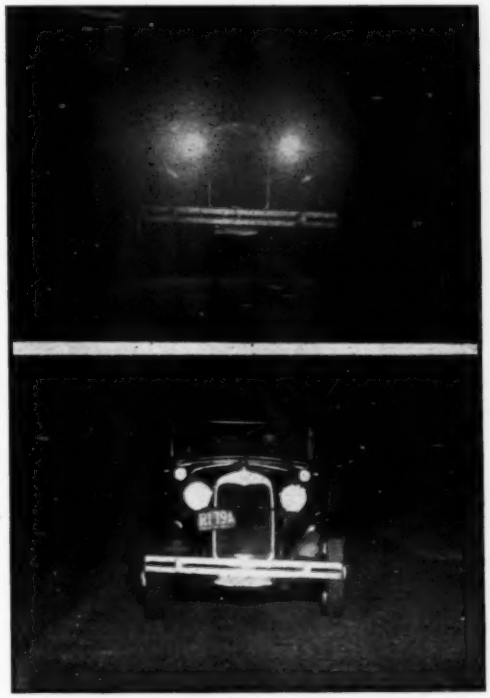


FIG. 9. Night photographs of a motor car with polarized headlights, illuminated by the light of another motor car beside the camera, taken without and with a suitably oriented polarizing filter over the camera lens.

using three sheets of polarizing material, two being rotatable.

When an ensemble of superposed sheets of Polaroid is used to control the relative intensity of light incident on a photo-cell or any other non-visual receiver, it is necessary to keep in mind the range of sensitivity of the receiver and to restrict the light to such wave-lengths for which Polaroid is a sufficiently good polarizer.

### Adjustable Sunglasses

The control of the relative intensity of visible light transmitted by two superposed sheets of Polaroid may obviously also be applied to sunglasses whose visual density is variable.[18]

### Light Valves

When intense electric fields are applied to certain liquids (nitrobenzene for instance), the liquids instantaneously become birefringent. This is the Kerr effect. When a Kerr cell is placed between crossed polarizers, the application of a transverse electric field permits light to traverse the system, its relative intensity being a function of the strength of the applied field. The availability of the new polarizers and their convenient form will undoubtedly lead to some common applications of the Kerr effect.

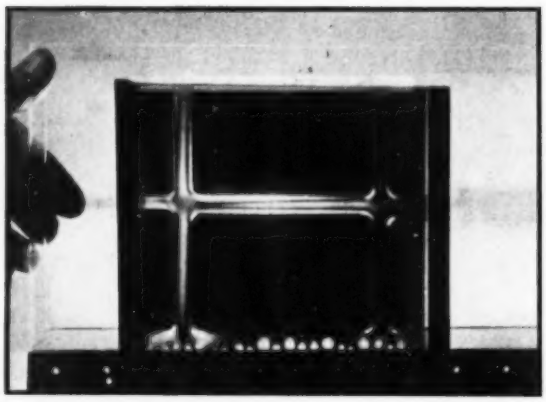


Fig. 10. Photograph of a "strain pattern" in a Bakelite model of a structural frame subjected to stress.

It might also be noted here that the new dichroic polarizers are widely being used in research laboratories for the identification of polarized spectrum lines in the Raman, Zeeman, and Stark effects.

### Optical Activity

When certain substances (quartz, turpentine, sugar solutions, etc.) are placed in the path of a beam of plane polarized light, they exhibit the ability to rotate the plane of polarization of the beam. At a given wave-length, the rotation per unit length of path is often a significant characteristic. For instance, it may serve as the measure of the concentration of sugar in solution. The high cost of polarimeters made with Nicol prisms has seriously restricted their use in college and industrial laboratories. Now it is possible to construct serviceable polarimeters at low cost, using simple half-shade devices of one sort or another.

## *III. Use of interference of polarized light*

### Colored Illumination—Advertising Displays

When a sheet of birefringent material (mica, Cellophane, cellulose tape, etc.) is inserted between two polarizing sheets, the ensemble becomes a color filter whose color may be varied by rotating one of the polarizing sheets or by tipping or changing the thickness of the birefringent sheet. This is an important new method of producing color in general illumination and advertising displays, and it may even find architectural applications, if the current trend toward the use of translucent building materials persists.[19]

The explanation of these color effects can be formulated in one of several ways. For anyone who is not a specialist in optics, the following explanation is perhaps as simple as any: The inclination of the polarizing axis of the first sheet of Polaroid governs the relative intensities of the two components of polarized light in the birefringent sheet. At the first surface of the birefringent sheet, corresponding elements of these two components are coherent and in phase with each other. But, since they proceed at different wave velocities, they develop an increasing difference of phase proportional to the thickness of the birefringent sheet. On emerging from the latter, they contribute two coplanar components to the plane polarized light transmitted by the second sheet of Polaroid; and these components are out of phase with each other. Under appropriate conditions (wave-length and thickness),

these contributions may be equal in magnitude and exactly a half-period out of phase with each other. They then annul each other by interference; and light of this wave-length will be subtracted from the original white light. At neighboring wave-lengths, only a part of the light will be lost by this process. The color of the transmitted light is the complementary to that lost by interference, and this is verified by the change of this colored residue to its complementary when either of the polarizing sheets is rotated through 90°. Needless to say, the colors vanish when the axis of either polarizer coincides with either of the principal directions of vibration of the birefringent sheet.

Interesting effects may be obtained by laying out designs in Cellophane, using different thicknesses and orientations, on glass and inserting them between polarizing sheets. Transparent cellulose tape (mending tape) is equally useful for this purpose. One of the polarizing sheets may be arranged to rotate slowly, driven by a motor, in order to change the color effects continuously.

## Detection of Strains in Transparent Materials

Whenever ordinary transparent materials (glass or certain plastics) are subjected to stresses, the strained regions inside are characterized by changes in their optical properties whereby they become temporarily birefringent. The amount of birefringence produced at any point in the material is in part a measure of the strains existing there. Consequently, when a strained sample is inserted between crossed polarizers, light is restored in the regions under strain. This restored light usually appears in the form of variously colored striations; and from the shape, spacing, and color of these striations, it is possible to identify the distribution of stress in the sample. This is the technique of photoelasticity.

This technique has long been used, albeit none too extensively, in the analysis of transparent models of structures, machine parts, etc. It has also been used for the identification of residual strains in glass products of all kinds arising from imperfect annealing.

Fig. 11. An industrial strain tester in use. The observer wears a pair of Polaroid spectacles which serve as the "analyzer."

The construction of photoelastic polariscopes with Nicol prisms was invariably encumbered with expensive lenses to make it possible to cover large fields of view with small Nicols. These encumbrances are now no longer necessary, and this fact is stimulating the introduction of photoelastic methods of testing and inspection into many industries, and it has revived interest in photoelastic research. A Polaroid-equipped photoelastic polariscope has been described by Tuzi and Nisida.[20]

## Polarizing Microscopes

With simple attachments, any microscope can now be converted into a polarizing instrument in a few moments, adapting it to the study of crystals, textile fibers, rock sections, etc. This ready availability of a polarizing microscope is rapidly leading to new developments in industrial testing and research. For instance, in the textile industry, as has been described by Schwarz,[21] the polarizing microscope can be used to establish the maturity of cotton, as well as to reveal variations in thickness of natural fibers.

Another interesting development is the fact that toy polarizing microscopes and polariscopes, with all the necessary accessories, are now available for the entertainment and instruction of children. A few years ago this would have seemed fantastic.

## *IV. Discrimination between specular and diffuse reflection*

### Photography

The use of polarizing screens and filters opens an entirely new technique in photography, and a part of this technique has already been mentioned in Section I. By using polarized illumination, in addition to a polarizing filter over the camera lens, the photographer gains a more complete control over the highlights in the picture. Therewith he also has the possibility of balancing contrasts.

In color photography, this arrangement makes it possible to control (and to increase) the saturations of colors. The light reflected from a colored surface may roughly be thought to consist of colored light scattered from the pigments and more or less white light reflected from the surface. By controlling the relative intensity of the specularly reflected light, the dilution of color identified with it is also under control.

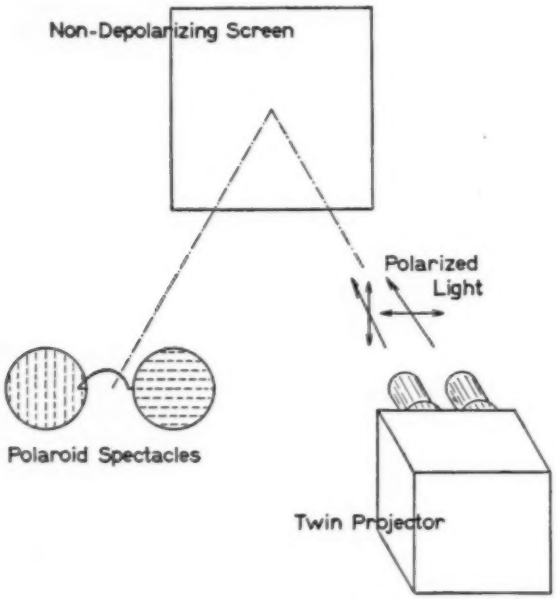



Fig. 12. Schematic diagram illustrating the projection of stereoscopic pictures with polarized light.

### Special Lighting Problems

In many types of work, where close attention to detail is imperative, the problem of glare is often very serious. Except where metallic reflection is involved, the glare can be reduced and often eliminated by using polarized illumination and polarizing spectacles or viewing screens. The plane of polarization of the incident light is turned until the polarizing axis of the viewing screen is crossed with the predominant direction of vibration in the reflected light reaching the eye.

This technique also facilitates the study of the texture of surfaces. In one of the Boston hospitals, a distinguished diagnostician regularly uses a Polaroid-equipped cystoscope with a small camera attachment. This instrument obliterates the wetness of tissue and makes it possible to study lesions with greater assurance than is otherwise possible.

Ophthalmoscopes equipped with polarizing filters are already on the market. These enable the oculist to view the fundus of the eye without being distracted by the glare of light reflected from the cornea.

Another interesting application was devised for a color printing establishment which wished to save the time necessary to allow a colored proof to dry before it could be compared with a standard chart. The problem was solved by a simple arrangement in which the two colored samples, side by side, are illuminated with polarized light and viewed through a Polaroid-equipped telescope of short focal length and low magnifying power.

## *V. Use in binocular vision*

### Projection of Stereoscopic Pictures

The impressions of three-dimensional depth in vision arise from the fact that we see simultaneously with two eyes side by side. One eye always sees a slightly different view than the other. Through lifelong experience, the mind is trained to interpret the disparity between the two ocular images in terms of a three-dimensionality of the outside world. Unless two objects at different distances are both so far away that their ocular images are practically identical, the mind immediately records which of the two objects is the nearer.

In order to project stereoscopic pictures on a screen, it is only necessary that there shall be a stereoscopic pair of pictures, one to be seen by the right eye and the other by the left. With polarized light this is accomplished very simply.

The projection of stereoscopic pictures with polarized light is schematically illustrated in Fig. 12. The two pictures are projected on a non-depolarizing screen, the light of the one being polarized at right angles to that of the other. The viewer wears polarizing spectacles, with the axes of the two polarizing "lenses" appropriately oriented and at right angles to each other to admit to each eye the image intended for it. Many of the details of stereoscopic projection with polarized light have been discussed by Kennedy[22] and by von Ardenne.[23]

Ordinarily this would imply taking two pictures with two cameras, side by side, and projecting them with a similar pair of projectors (linked together in the case of moving pictures). This inconvenience may however be avoided with the aid of a patented device which can be attached either to the camera or to a projector. The operation of this device involves two contiguous images on a single frame of film.

This method is the only one known to project stereoscopic pictures in natural colors. It may also be adapted to the construction of stationary viewing devices for stereoscopic pictures in museums, schools, etc.

### Ophthalmic Instruments

The newest developments in ophthalmic optics indicate that the refraction of the two eyes relative to each other is fully as important as the correction of the vision of individual eyes. Many mysterious cases of defective vision have been found to involve a disparity of magnification of the two ocular images. This defect of vision was first studied by Ames and Gliddon[24] and is now known as aniseikonia. The study of cases involving this defect is now made more easily than before with the aid of polarizing screens to give the eyes independent images.

### Experimental Psychology

The problems of binocular vision are gradually becoming important in experimental psychology, especially in connection with color vision. These[25] and others may now be studied with the aid of polarizing filters to keep the ocular images independent of each other.

Many other applications of polarized light are known and in use, and it is to be expected that more will be devised in the future. It is hoped that this summary may serve to stimulate such new developments.

It should also be noted that this list is restricted to applications of plane polarized light. The complex types of polarization (circular and elliptic) are equally amenable to practical uses, particularly since synthetic birefringent sheetings are now commonly available.

### *Bibliography*


1. A. M. Marks, U. S. Patent 2,104,949.
2. F. Bernauer, Fortschr. d. Min. **19**, 22 (1935).
3. E. H. Land, U. S. Patents 1,918,848; 1,989,371; 1,951,664; 1,956,867; 2,011,553; etc. Also foreign patents.
4. W. B. Herapath, Phil. Mag. **3**, 161 (1852).
5. C. D. West, Am. Mineralogist **22**, 731 (1937).
6. S. Roesch, Fortschr. d. Min. **21**, 89 (1937).
7. A. Zimmern and M. Coutin, Comptes rendus **182**, 1214 (1926).
8. Pollard, Nature **138**, 311 (1936).
9. H. Freundlich, Chemistry and Industry 698, July 31 (1937).
10. M. Grabau, J. Opt. Soc. Am. **27**, 420 (1937).
11. M. Haase, Zeiss-Nachrichten, 2. Folge, Heft, 2 August, 1936.
12. F. Jentsch, Zeits. f. tech. Physik **7**, 310 (1926). G. J. Pokrowski, Zeits. f. Physik **30**, 66 (1924). Woronkoff and Pokrowski, Zeits. f. Physik **30**, 139 (1924).
13. Malus, Mem. de Phys. et de Chim. de la Soc. D'Arcueil **2**, 143 (1809). This paper is reprinted in the appendix of E. Mach, "The Principles of Physical Optics."
14. Hulburt, J. Opt. Soc. Am. **26**, 216 (1936).
15. Tuttle and McFarlane, J. Soc. Mot. Pict. Engrs. **25**, 69 (1935).
16. B. Alfieri, Brit. J. Phot. Alm. 1937, p. 159.
17. L. W. Chubb, U. S. Patent 2,087,795, Trans. Ill. Eng. Soc. **32**, 505 (1937); E. H. Land, J. Frank. Inst. **224**, 269 (1937).
18. E. H. Land, U. S. Patent 2,005,426.
19. E. H. Land, U. S. Patent 2,018,214.
20. Tuzi and Nisida, Sci. Pap. Inst. Phys. & Chem. Tokyo **31**, 99 (1937).
21. E. R. Schwarz, J. Opt. Soc. Am. **27**, 44 (1937).
22. C. Kennedy, J. Soc. Mot. Pict. Engrs. **26**, 3 (1936).
23. M. v. Ardenne, Zeits. f. tech. Physik **17**, 332 (1936).
24. Ames and Gliddon, Trans. Ophth. Sec. Am. Med. Assn. 1928.
25. E. H. Land and W. A. Hunt, Science **83**, 309 (1936).

# Frederic Eugene Ives

By Edward Epstean and John A. Tennant

## I

IT is of interest to note that while many men of scientific training have contributed largely to the invention and perfecting of the photographic and photomechanical processes, a very considerable part of the advancement made since the beginning of photography must be credited to the original research and discoveries of men who lacked a specialized scientific training, and who acquired their insight and understanding of the scientific principles involved in the solution of their problems by self-study and experience.

Wedgwood, Sir Humphrey Davy, Fox Talbot, Sir John Herschel, Robert Hunt, Clerk Maxwell, Wollaston, Petzval, Draper, Carey Lea, Eder, Vogel, Becquerel, Marey, August and Louis Lumière, Abney, Hurter and Driffield, Mees, Sheppard and Clark are familiar names among the scientists to whom we are indebted for notable advances in photography. On the other hand, the untrained, amateur scientists or "born inventors" are very numerous, and the quality of their achievement is in no way inferior to that of the more favored scientists. They came into photography from a diversity of occupations: Niepce was a cleric and soldier; Daguerre, a painter of dioramas; Bayard, clerk in a French Ministry; Niepce de Saint Victor, an officer in the army of Napoleon; Reade and Hannibal Goodwin were clergymen; Maddox, a medical man; Morse, a painter; Scott Archer, a sculptor, and so on.

It is with one of these untrained scientists, *Frederic Eugene Ives* that these biographical notes are concerned. Born on a small farm near Litchfield, Connecticut, February 17, 1856, he traced his ancestry to English families of remote Norman-French origin, who came to New England about 1635. A shy and solitary boy, he was given to dreaming and wandering about the fields and woods; with a vivid sense of the glory of color in the world about him, and a predominant eye-mindedness by which he could visualize the relationship of material things as though they were actually before his eyes.

His schooling, over brief periods and in different places, hardly carried him through primary arithmetic. It ended abruptly before he was twelve years old, on the death of his father, who had abandoned the farm and taken over a general store at Norfolk, Connecticut. About this time the boy had acquired a "dollar microscope" and found an old school textbook on natural philosophy. These opened up to him a new world, awakening an interest in light and optics that became almost a passion—and the fruitful source of inventions in later years.

Facing the problem of earning a living, before he was fourteen years old, young Ives apprenticed himself for three years (1870–72) in the printing office of the Litchfield *Enquirer*, and so became a journeyman printer in his seventeenth year.

It was apparently in the second year of his apprenticeship that he made his first experiments in photography, with the wet-collodion process then in general use, and a camera which he made from a cigar box, fitted with a spectacle lens.

He also printed and published, through 1871–72, independently of his work on the *Enquirer*, a small "amateur paper:" *The Lilliputian Advertiser*, with a mixed content of poetry, a serial story, editorials, pithy paragraphs—"wise and otherwise," and a column or two of card advertisements as a source of revenue.

It was for the purpose of occasionally illustrating this little venture in journalism that, early in 1872, he purchased some engravers' tools and taught himself wood engraving. And it was the painful slowness of the wood engraver's hand work in cutting a line reproduction of a photograph, for the July 1872 issue of the *Advertiser*,

that turned his thought to the possibility of making printing plates by photography, and led him to work out the photomechanical processes which illustrate our newspapers, periodicals and books today.

The term of his apprenticeship in the printing office of the *Enquirer* ended early in 1873, and he worked as a journeyman printer in Ithaca, New York, for the rest of the year. There are entries in his *Diary* of 1873 which tell of attendance at the lectures on light and spectrum analysis, given in April of that year by Professor William A. Anthony, of the department of physics at Cornell University. Other entries in the *Diaries* of 1872 and 1873 speak of continued photographic experiments and improved equipment. In 1874–75 he broadened his knowledge of practical photography by working in professional studios, thus gaining familiarity with its commercial and industrial applications.

It was on the strength of this experience that, late in 1875, he was engaged to do the photographic work of the physics laboratory at Cornell University, under Professor Anthony. There, with access to the technical library and lectures of the university, he found time and means to develop his ideas on the photomechanical production of printing plates.

During the years at Cornell, also, inspired perhaps by the researches of Vogel and Becquerel, then recently published, he began the experiments in correct color-tone photography (the ortho- and panchromatism of today) and color reproduction, which resulted, some years later, in his system of composite heliochromy, by which the colors of nature were perfectly reproduced to the eye in photographic images.

In these two fields of applied photography we find the outstanding achievements of his life—with, as will be seen later, many other inventions of merit and usefulness.

When Ives, as a youth of twenty, began his work in photomechanical reproduction processes, the only method of preparing printing plates (illustrations) that could be incorporated with the text (types) in the forms on typographic presses was wood engraving—a stereotype of the engraved wood block being used in commercial practice.

## II

The first of the Ives engraving processes, perfected at Cornell in 1877, was his swelled gelatine-plaster-cast-stereotype process of line photoengraving. As the hyphenated name indicates, the process produced a stereotype printing plate molded from a plaster cast of a swelled gelatine relief image of the design photographed. This relief, which replaced the laborious hand and tool work of the wood engraver, was obtained by exposing a film of bichromated gelatine to light action in contact with a photographic negative of the design, after which the film was allowed to swell in water into relief.

This photo-stereotype process proved itself, in two years' use in illustrating the college paper *Cocagne* and other demonstration work, to be a thoroughly practical and serviceable method of producing line illustrations for printing on typographic presses. With the basically similar

FIG. 1. Frederic E. Ives in his Philadelphia laboratory, about 1930.

processes introduced about the same time by Moss, of New York; Mumler, of Boston, and Louis Levy, of Philadelphia, it obviously stemmed from the pioneer work in gelatine relief engraving of Paul Pretsch in Austria, about 1854. It marked the beginning of the transition from wood engraving to photoengraving, and continued in general use until the introduction of the typographic relief-etching processes of today.

While operating his line photoengraving method at Cornell, and busied with the daily stint of general photographic work for the university, Ives now turned his thought and experimenting to the more difficult problem of "halftone" photoengraving. This meant a photographic method of producing printing plates direct from photographs, pencil and wash drawings, and similar "continuous tone" pictures. In these, as distinguished from pictures in simple black and white, the light and shade effects of the subject are represented by smooth shadings or "halftones" between the deepest darks (blacks) and the highest lights (whites) of the picture.

It was a problem that gave him many a headache. As a practical printer he realized that the inking method employed on the typographic press necessitated the breaking up of these smooth shaded areas into a pattern of definite lines and dots, and allowing the white paper to show between, or rather covering the paper by graduating the size of the lines or dots according to the tone values of the different shade areas in the original. But how to do it seemed, for a long time, past finding out.

## III

In the modern halftone process, this breaking up of the ink surface of the plate for the reproduction of the fine detail and halftones, and the graduation of lines and dots to approximate the tone values of the original, is readily accomplished by the proper use of an optical device known as a halftone screen or grating. This consists of two cemented glass plates, each covered with fine opaque lines of a width equal to the transparent space between them, and usually ruled diagonally to the horizontal and vertical directions of the plate. These plates, when sealed face to face so that the lines cross, form a network or grating which, placed at a calculated distance in front of the sensitive plate in the camera, translates the smooth shading of the photographic image into a pattern of definitely black lines and dots, graduated in size according to the intensity of the light reaching the plate through the tiny transparent apertures of the screen, controlled by a determined relationship between the lens aperture used, the distance separating the screen and sensitive plate, and the exposure.

This remarkable tool, suggested by Fox Talbot in connection with his photogravure process of 1852, embodies in its present state, the research of hundreds of workers in photomechanical reproduction over almost fifty years (in which Ives had a notable share), and was perfected in its present form by Louis and Max Levy, of Philadelphia, about 1890.

When Ives faced his problem in 1878, however, the halftone screen was still in a crudely experimental stage and little known. His thinking and experimenting were therefore based upon his experience in wood engraving. In this art, as he knew, the tones of the original being engraved are simulated by varying the width of the lines by the skillful use of a V-shaped cutting tool which, in proportion to its penetration of the wood block produces lines of varying surface width.

It was on this principle, remembered from his early attempts at wood engraving, that Ives worked out his first "halftone" photoengraving process, in the summer of 1878, at Cornell. His description of it follows, duly witnessed and dated August 12, 1878, on which day he had produced conclusive results with the method.

"1st. From an ordinary photographic negative a relief in gelatine, similar to that used in the Woodburytype process, but perhaps in lower relief, is obtained.

"2nd. This relief is carefully and uniformly inked, with fine printers' ink, and pressed between two flat surfaces or between rollers, against paper or other material upon which is stamped or otherwise produced a fine grain or other suitable surface. The inked relief being highest in the black parts, presses down the grain of the paper on the corresponding parts, and the removal of

the ink, by the paper, from those parts of the relief, produces a black impression, while upon those parts where the relief of the gelatine is lower, the grained surface is pressed less, and the ink taken up in spots the size of which depends on the grain of the paper and the amount of pressure, and producing an effect similar to that of crayon sketches upon such a surface.

"3rd. Relief plates may be made from this impression either by the usual photo-typographic processes or, perhaps, by obtaining a cast or electrotype of the impressed paper or other material used to receive the impression from the gelatine relief."

It should be noted here that the "paper or other material upon which is stamped or otherwise produced a fine grain or other suitable surface," mentioned in the second clause of the statement, was, in fact, paper impressed with V-shaped grooves giving a relief ruling of about 25 fine lines to the inch. This, as Ives later disclosed, was, in effect, his application of the relation of an engraver's V tool to the varying heights of the photographic relief, which produced the necessary graduation of lines (and dots, by crossing the lines)—the fundamental principle of the "Ives Process." Later, Ives substituted an optical V-line for this mechanical V-line, by the use of a crossline sealed screen, as in today's practice.

Circumstances prevented further development of the process at that time. The first prints from finished plates by this method, made in Philadelphia in June 1879 and examined by the present writer in the Ives *Scrap Books* at the Franklin Institute, are extremely interesting as proofs of an unperfected process of that remote date.

By a strange coincidence, Charles Petit, of Paris, patented (in France) August 1878, a halftone process very similar to that invented by Ives a few days earlier, embodying the same mechanical principle in a slightly different procedure. When Petit, a year or so later, sought to obtain a patent for his process in the United States, Ives was able, by means of the statement given above, to establish his claim of earlier invention, and judgment of priority was awarded to Ives, sustaining his patent of February 8, 1881.

Meanwhile, in October 1878, lacking means to patent and develop his process commercially, Ives left Cornell and went to Baltimore to make line engravings for an illustrated newspaper, which was abandoned by its owner after a few months.

## IV

In February 1879 Ives went to Philadelphia, to establish a photoengraving department for the Crosscup and West Engraving Company, wood engravers of repute, who "believed in the commercial importance of the invention (line photoengraving), and were particularly impressed by the promise of the "halftone" process, but had no money to invest." With this somewhat dubious encouragement the youthful inventor, supplying the necessary camera and equipment, entered the employ of this firm and remained there during the ten or twelve years following. His *Diary* of 1879, under the date—Saturday, March 1—has the laconic entry: "Received first week's pay ($15) from C. & W."

Those early years in Philadelphia were so crowded with activities in photographic and photomechanical research, and inventions in the field of optics and related sciences, that it is difficult to preserve even a semblance of chronological order in this brief record of them. Perhaps the "incessant urge to do useful work," which, we are told in his *Pioneering Inventions*, he inherited from his father, who "practically worked himself to death," may account for this untiring effort.

Among these activities, with the labor, at first single-handed, involved in making the new business of line photoengraving profitable to his employers, the perfecting of his halftone process occupied the first place. The record speaks of continual experimenting and improvements in this process. A white surfaced plaster cast of a swelled gelatine relief was substituted for the Woodburytype relief print, and on this an inked surface of V-shaped lines and dots was impressed with an elastic rubber sheet covered with pyramidal raised points and lines. The inked impression on the plaster cast was then used, like an artist's pen drawing, to produce a photo-typographic printing plate.

This improved process, patented in 1881, was commercially introduced early in that year, and found great favor among the publishers of magazines and books. The *Century*, *Harpers' Magazine*, *Godey's Lady's Book* and other peri-

FIG. 2. Enlargement of early halftone, showing rendering of tone values by variable sized printing surfaces.

odicals of the 'eighties contain many examples of this "Ives Process," printed alongside the work of famous wood engravers of that time. Prints from "Ives Process" blocks appeared in the technical journals here and abroad, and were everywhere praised as a distinct advance in the art of illustration. As so perfected, this "Ives Process" was employed in all the commercial work done by Crosscup and West until Ives introduced his "screen" halftone process in 1885–86.

Another improvement or variation of the "Ives Process" apparently worked out in 1883, was a method of collodionizing the inked plaster impression, stripping the collodion film and making a negative from it by contact printing, from which transfers to stone and metal were made for lithographic printing. The proofs or examples of this method, found in the *Scrap Books*, are equal in quality to the work of this class at that period, but we have no information of its commercial use.

The photoengraving business of Crosscup and West had, by this time, increased to such an extent that, in February 1881, an agreement was made with Ives, by which they secured the exclusive right to operate the Ives processes with later developments in this field, as well as to license others for the same purpose. It is characteristic of Ives that this agreement reveals no material advantage accruing to him, beyond increased facilities for work and experiment.

## V

We get another glimpse of his genius for invention in an article on "Photo-Telegraphy" in the July, 1881, issue of *The Philadelphia Photographer*. In this article, noting the lack of success in attempts then being made to transmit photographic images by telegraph, in which a selenium cell was used in an electric current, Ives suggested the substitution of a purely mechanical device in place of the selenium cell, for the purpose of controlling the electric current by the aid of a photographic relief.

The principle here disclosed by Ives was later used by Eaton, Amstutz and Belin in their phototelegraphic experiments. In a later note Ives pointed out that the device could be utilized to produce a "halftone dot" formation of the image, corresponding to the use of an intermediate halftone process plate, which was patented many years later.

In that same year also, Ives invented an ingenious attachment for the magic lantern, described as an "ether saturator." This was a small, portable device by which the vapor of ether was safely and conveniently substituted for hydrogen or coal gas in producing the oxyhydrogen blowpipe flame and lime light then commonly used as the illuminant in lantern slide projection. The Franklin Institute awarded a medal to the inventor for the efficiency and usefulness of this apparatus, in connection with a portable (folding) projection lantern invented a few years later. It was used for several years in connection with the illustrated lectures of the institute and elsewhere, until the electric current became generally available and superseded all other forms of illuminants for projection purposes.

Mention has already been made of experiments in correct color-tone photography made by Ives while at Cornell, in 1878. An account of these experiments was published in the December, 1879 issue of *The Philadelphia Photographer*, but received little or no attention at that time. In 1881, however, Ives applied the "panchromatic" method derived from his experiments, with his halftone method of 1880, to the production of trichromatic halftone plates and prints. This was pioneering "with a vengeance"—the earliest attempt to link color photography to the printing press. The results, time and circumstances being considered, were excellent, but the significance of the experiment escaped notice until the public showing of one of the prints (with the original chromo-lithograph alongside) at the Novelties Exhibition in Philadelphia, in 1885, aroused much interest and favorable comment. Later Ives improved this "panchromatic" method and applied it in the production of lantern transparencies in color, which he showed at the Franklin Institute in 1885–86. For this work he was awarded the Franklin Institute and John Scott Legacy medals in the course of these two years.

In the midst of these diversions, and an ever increasing volume of commercial work coming in as the excellence of the "Ives Process" became known, the restless spirit of Frederic Ives began to chafe under the limitations of the mechanical V and the swelled gelatine relief in the reproduction of halftones.

In that year a chance meeting with Stephen Horgan, who, in 1880, had transferred to stone screen halftones made direct in the camera and printed them (lithographically) in a New York newspaper, may have shown him the possibility of translating tones into lines and dots by means of a line screen placed out-of-focus. Perhaps this suggested to Ives that he could make light do the work of the engravers' V tool. We do not know.

In 1882 Meisenbach, of Munich had scored a success in the photomechanical world with a screen process, in which he obtained the graduation of lines by a slight right-angled movement of a single line screen during the exposure—as Swan had done some years before. Perhaps this idea intensified the mind of Ives to the possibility of an "optical V-line."

It is known that in 1882–83, Ives was thinking of an "optical V" to replace the mechanical V of his earlier process. That he linked the use of a screen to his thinking is seen in an entry in his *Diary* of 1883, under date June 26, which reads: "Made a halftone plate direct from a photograph by using a grating properly (Un)focused to produce graduated lines, etc." Other entries during 1883 show that he continued his experiments with screens, or gratings as they were then called.

As he was an active member of the Franklin Institute and an omnivorous reader, these experiments were doubtless accompanied by wide reading in the foreign patents and technical periodicals files of the institute. In the later papers on his inventions there are many references to the attempts made by others in the design and use of screens in photomechanical methods. The general failure of these attempts was, in his opinion, due to a "comparative if not complete ignorance of the mechanical and optical principles involved" in the problem of translating the body shades of tone subjects into lines or dots of definitely graduated surface areas reproducing, in effect, the smooth shading of the original.

His reading and his experience strengthened his conviction that the fundamental principle of success in these methods was in utilizing the relation of an engravers' V tool to the photographic image, which Petit and he had used in their photoengraving methods.

We must remember that in these new experiments with screens, Ives was seeking, not only to substitute an optical V for the earlier mechanical V, but also a simpler and less roundabout method —wholly photographic—of making halftone negatives direct in the camera.

His first experiments were made with a single line screen, placed out of focus in relation to the image on the sensitive plate, and a diamond-shaped stop in the lens. This gave him a graduated line as expected, but did not satisfy. These experiments continued during 1884–85, with constant variation in the application of the factors by which he sought to control the graduation of the lines.

## VI

Before October 1885, having exhausted the possibilities of the single line screen in this direction, Ives conceived the idea of sealing two single line screens, with equal black and white spaces, face to face. The minute transparent apertures of this screen, as calculated by Ives from a theoretical postulation of its behavior, functioned like pinhole cameras to produce dot images of the lens aperture, brightest at the center, but shaded regularly towards the image of the adjacent (screen) aperture. He had achieved the substitution of an optical V-line for the mechanical V-line, and the basis of a direct method—wholly photographic—of obtaining halftone negatives in the camera, without the tedious after work of the gelatine relief method. An entry in his *Diary* of 1885, dated Monday, October 5, reads: "Made first halftone job for C. & W. by new quick method. (Screen process.)"

In practice with crossline screens, graduated dots are substituted for graduated lines, these effecting a better rendering of fine details and the more delicate shadings. The qualities of graduation are regulated by screen apertures, distance of screen from sensitive plate, size and shape of the lens aperture, and the relation of these to each other. It was by a carefully calculated system of the adjustment of these factors that Ives perfected, in 1886, his screen halftone process. With minor changes in procedure, this system is still in universal use, after more than half a century.

On the advice of his employers, Ives did not patent this process and it was operated as a trade secret for a few years. The optical V-line and the scientific principles involved in its production and use were cautiously explained in a lecture at the Franklin Institute, and published in the May 1888 issue of the *Journal* of the institute, but its application with the crossline screen was not fully disclosed until Ives described the theory and practice of the process, in a series of lectures delivered at the Bolt Court Technical School, London, in 1898.

Between these years (1888 and 1898) many things happened. The Ives screen process had replaced the relief method since 1886, at first occasionally, later altogether, with the coming in of the baked enamel copper-etching method, which was worked out in Ives's workroom in the Crosscup and West establishment in 1886. The general adoption of the Ives process, as it became known by leakage of details through the disloyalty of employes, was greatly promoted by the commercial introduction of the perfected crossline sealed screen by Max Levy, about 1890, and, later, the introduction, by Levy, of the acid-blast etching machine.

During the last few years (1885–1891) of his engagement with Crosscup and West, Ives gave much of his attention to his isochromatic (color correct) method and color reproduction. He developed the early isochromatic photography with chlorophyl process of 1878, now combining eosine with chlorophyl, and thus extended the color-sensitivity of the photographic plate through the whole of the visible spectrum. His work with these color-sensitizers was summarized and illustrated with color charts, photographed with and without color-sensitive plates and color screens, in papers read at the Franklin Institute and published in the institute's *Journal* 1885–1886. Incidentally the use of these color-sensitive plates necessitated the preparation and first use of the bichromate cell; glass sealed, dyed gelatine color screens, and the green darkroom safelights now commonly employed with color-sensitive emulsions. And it was with negatives made by these methods that he made a convincing demonstration of the additive trichromatic process at the Franklin Institute in February 1888—a reproduction by triple lantern projection of an autumn landscape in colors.

About this time, also, he repeated and described certain experiments first made at Cornell, in 1878, in photographing by the aid of the phosphorescent tablet. In renewing these experiments he discovered a means of photographing small obscure metal objects by the action of heat radiations. One of the "heat photographs" thus obtained—an impression of three pins on a solarized phosphorescent tablet of the limelight spectrum—he thought to be one of the earliest known heat photographs. The papers on this subject were published in the *Journal of The Franklin Institute*, January and March 1887.

An important paper on "Correct Color-Tone Photography With Ordinary Gelatine-Bromide

Plates" was read before the Franklin Institute in June 1886 and published in the August issue of the *Journal*. In this paper Ives tells of the preparation of the compound light filters (mixtures of aniline color solutions) which proved so effective in his heliochromoscope of 1891.

These experiments in correct color-tone photography—the correct rendering by photography of the tone values of the color luminosities of objects—were interesting in themselves as recording scientific investigations in a comparatively unexplored field, and extremely valuable as applied in everyday photographic work, by means of the practical methods which he freely published in the technical journals of the time. To Frederic Ives, however, they were chiefly interesting as the foundation of his most notable achievement in photography.

This was his system of composite heliochromy —the photographic representation of colored objects in their natural colors, as seen by the eye—based on the Young-Helmholtz theory of color vision as demonstrated by Clerk Maxwell in 1861, according to which the eye has three color sensations: red, green and blue-violet, all colors being perceived as mixtures of these three primary colors to a different extent. Ives first described his system in a lecture at the Franklin Institute in February, published in the May 1888 issue of the *Journal* of the Institute, and in a more detailed statement in the *Journal* of January 1891, as embodying a "new principle in heliochromy."

Briefly summarized, this new principle is that of making sets of negatives by the action of light rays in proportion as they excite primary color sensations, and images or prints from such negatives with colors that represent primary color sensations. The system, based on Maxwell's photometric measurement of the effect of the spectrum upon these primary color sensations, involves the production of one negative by the joint action of the red, orange, yellow and yellow-green rays, in definite proportions, to represent the red sensation; one by the joint action of the orange, yellow, green and green-blue rays, in definite proportions, to represent the green sensation; and one by the joint action of the green-blue, blue and violet rays, in definite proportions, to represent the blue sensation.

These three negatives are produced in a single camera, with a single lens, by simultaneous exposure on a single plate, by means of color screens of compound composition, the three images so obtained reproducing the light and shade of the object photographed as they appeared to the respective fundamental color sensations, these images being thrown, by means of mirrors, on to different parts of the sensitive plate.

From this triple negative a triple positive is made, and the three images being exactly superposed by projection in a triple projection lantern, through filters or colored screens: one of red, one of green, and one of blue-violet, the combination gives a faithful reproduction of the object as seen by the eye. This was the way in which Ives exhibited the first example of his method. For the more convenient viewing of the color reproductions Ives constructed an ingenious instrument, variously called heliochromoscope, photochromoscope or kromskop, in which the three positives were placed over colored glasses and their images exactly superposed by means of mirrors or reflectors. Lantern slides were also made from the triple negative, the three colored positives (produced by a dye imbibition method) being cemented together in exact register for use in the ordinary, single optical lantern. Color prints on paper could also be made from the triple negative by pigment printing (trichrome carbon, etc.) or dye toning. These methods were, however, not as simple or as fully developed then as now.

The principal optical devices and mechanical aids employed in the carrying out of composite heliochromy were designed by Ives as called for in the development of the process: special trichromatic cameras, triple projection lanterns, and viewing instruments (photochromoscopes) in many different models. The process was patented in 1890; improvements in this and the apparatus mentioned were similarly protected in the years 1892–1908. Among the later applications of the process were diffraction color photography, with special forms of the photochromoscope for this use; a stereoscopic photochromoscope, which added the illusion of relief to the color reproductions; and a "miniature" kromskop, of simple and inexpensive construction, in which the

three images were disposed in a line on a single plate.

In the method of composite heliochromy thus far given, the color reproduction is obtained by analysis and positive synthesis with three negatives. Ducos du Hauron, in his patent of 1868, which contains the germ of almost every method of color photography "discovered" since his day, suggested a method of obtaining analysis and synthesis in a single operation. He proposed to make a single negative through a screen made up of fine juxtaposed colored lines, alternately red, yellow and blue, and then to view a positive from this negative through the same screen, making the red lines of the screen cover the lines which were made in the negative through red lines, and so on. It is evident that, with lines fine enough not to be separately perceived by the eye, the result should be a color record requiring no viewing device to see it in its perfection. A brilliant conception, afterwards applied in the methods of Joly, McDonough, Warner-Powrie and other screen plate processes. In 1895 Ives thought of a method of making such negatives, permitting of a single exposure for all three color lines with an opaque line screen, and in 1899 devised a means of synthesis for such pictures without the colored line screen, and he applied for patents on both inventions—an application of his principle (1886) of screen "pinhole images" of the lens aperture in his halftone process.

About this time Ives worked out a method of producing "subtractive" process lantern slides—the first superposed dyed gelatine relief positives, for transparencies and paper prints. The application of his heliochromic method in negative making in the trichromatic halftone process was first mentioned in a paper read at the London Camera Club Conference of 1894, but was not perfected until 1901, and patented in 1902. This process was employed commercially, under the supervision of Ives, in New York, about 1904. The work turned out was of remarkably high quality, but lack of capital and "Union" opposition led to the abandonment of the enterprise a year later.

## VII

The announcement and demonstration of composite heliochromy by Ives, in the principal Eastern cities during 1888–1890 excited widespread interest. The newspapers devoted columns of small print with detailed descriptions of the process and praise of the inventor; the magazines printed pictures of "Ives at his photochromoscope;" and the public halls were all too small to accommodate the crowds desiring to see actual photographs in color, direct from nature.

The excitement crossed the Atlantic, and Ives was invited to read a paper on composite heliochromy at the Royal Institution, London, in May 1892. On his arrival in England in April of that year, his invention was hailed with an enthusiasm even greater than that in America. He had arranged for a visit of a couple of months, but his lecture engagements and demonstrations before the Royal Society, the Society of Arts and other scientific and photographic societies caused him to prolong his stay over a year. The perfection of his color reproductions, the ingenuity of his heliochromoscope and triple projection lantern, and his ability as a public speaker captivated the crowded audiences at his lectures, and Ives was acclaimed, even by conservative scientists of note, as having solved the problem of color photography on scientific principles. Sir William Abney, Sir Joseph Wilson Swan, Sir William Herschel, Silvanus Thompson, Dr. Eder, F. Stolze, Henry Trueman Wood, Thomas Bolas and other well-known authorities in the photographic world declared the work of Ives to be a distinct advance on all that had been accomplished up to that time.

During that first English visit, Ives was awarded the medals of the Society of Arts, London, the Vienna Photographic Society, and the Photographic Society of Philadelphia, and elected to Honorary Fellowship in several English and Continental scientific societies. Returning to the United States in April 1894, he set about the perfecting of the photochromoscope and other inventions.

## VIII

The years 1895–1898 found him again in Europe. This visit had a twofold purpose: to introduce his screen halftone process in the establishment of the Swan Engraving Company, London, and to effect the commercial introduction there of the now perfected photochromoscope.

The success of the earlier visit was repeated, and a syndicate formed in 1896 to exploit the kromskop, kromograms, and accessories of the Ives color processes. There were visits to Vienna and other continental cities, and lecture-demonstrations in many of the larger towns of England and Scotland. Wherever he went, Ives was given an enthusiastic welcome. The Scottish Society of Arts, Edinburgh, gave him the Keith medal ("valued at fifteen sovereigns, the cost of the inscription, usually about £1, being customarily deducted from the prize sum," as was explained to him). He lectured again before the Society of Arts, to report on the progress made in his color work; and before the Royal Photographic Society, where he was elected to the Council; the Camera Club, of which Sir William Abney was president, and many scientific and technical societies.

By this time, however, a number of other inventors, with other color processes, had come forward to share the popular interest awakened by Ives and his kromskop, and the ever fickle public turned to the exhibitions of color photography by Lippmann, Joly, Lumière, Chassagne, Veress, Bennetto and others. The Ives color reproductions, however perfect, could not be hung on a wall or passed around like a paper print. The making of kromograms for the kromskop demanded more skill than the average amateur could give to it. And there were unkind critics who suggested that the kromoskop, then coming into the shops, was a scientific toy, which at any moment might be replaced by a simple, direct method of color photography.

The British Kromskop Syndicate was liquidated in 1898 and Ives prepared for the return to America. Before leaving England, however, he gave a course of lectures on "Photo-Process Work," at the Bolt Court Technical School, London, in the early part of 1898. In these lectures, with illustration and practical demonstration, he set forth the history and development of photomechanical reproduction processes, stressing the theory and practice of his own original methods, with generous credit to other workers in the field, and giving freely from his wide experience for the helping of those who made the course. The lectures were afterwards published in the *London Education Gazette*, No. 51, January 1899.

After the feverish activities of his four years in Europe, Ives, doubtless glad to be free once more to develop new inventions, settled down to supervise the manufacture of trichromatic cameras, kromskops and accessories for the neglected home market. The *Scrap Books* of this period speak of the incorporation of the Ives Kromskop Company, of Philadelphia, with a capital stock of $1,000,000 and an imposing board of directors, but are silent as to its activities or end. It would seem that his experience as the head of this enterprise convinced Ives that the kromskop could not be made profitable. He moved to New York to take up the making of trichromatic printing plates. As already told, this adventure ended in failure.

In 1900 the Royal Photographic Society, London, showed its appreciation of the work Ives had done in the advancement of photography, by inviting him to give the Traill Taylor Memorial lecture of that year. He responded with a paper on "The Optics of Trichromatic Photography," in which he gave a concise and critical review of the history, science and mechanics of the indirect methods of pigmentary color photography. Apart from its fairness in discussing the methods suggested by others in this special field, the paper reveals, in a striking way, the clear perception of the scientific principles which guided Ives in his development of heliochromy.

The Science Medal of the Society was awarded to him in 1902, followed by the coveted distinction of the Progress Medal in 1903. In this year, also, the Franklin Institute recognized his outstanding work in heliochromy and photoengraving, by awarding to him the Elliott-Cresson Medal—the highest honor in the gift of the institute.

In nowise dismayed by the unlooked for collapse of the color plate engraving enterprise in New York, Ives now returned to Philadelphia and to his old vocation—invention. We recall the boy Ives as the proud possessor of a "dollar microscope." In middle life he became an ardent microscopist and Fellow of the American and Royal Microscopical Societies. From this interest came the invention of a photomicrographic camera of special design, patented in 1902, and

the modern short-tube, single objective, binocular microscope of 1903. A further contribution to this branch of science was a useful method of stereoscopic photomicrography with high powers, published in the *Transactions* of the American Microscopical Society, November 1903.

The Ives Parallax Stereogram, a device producing stereoscopic relief without the use of a stereoscope (1904) was one of the most popular of the inventions of this time, and found a wide usefulness as a business show window attraction. It was an application of the line screen in the production of a transparency of a design or photograph, which caused the flat design to stand out with the boldness of life and actuality.

Other specialties or inventions for scientific, educational and industrial use, which made their appearance in this period, were: diffraction grating replica; a diffraction photochromoscope (1906–07); a trichromatic colorimeter and tint photometer (1908–09); flexotone and tempoflex, two devices controlling tone expression and tempo in automatic pianos (1908–09); and a plastigmat system for the production of anaglyphs and viewing device, giving pictures in stereoscopic relief for theaters.

After 1910 Ives turned again to color photography, and devised several methods of making color records and color prints, for amateur photographers. Some of these, notably the tripak system (1912): a box hand camera and plateholder designed for the use of a "pack" of three plates, handled as a unit until after development; and hicrography, a similar system, but designed for professional portraiture; together with methods of producing color prints by dye imbibition, were popular for a year or two after their introduction, but failed in the absence of "finishing" or processing service plants which "do the rest" for the modern amateur.

The application of his wide knowledge and experience to cinematography in colors resulted (1913–1921) in numerous inventions looking to the production of two-color motion pictures on ordinary cine positive film. As in the development of composite heliochromy, so in the motion picture field, his activities included the designing and invention of special cameras and physical equipment as needed (1916–19); light splitting devices for cine color cameras, permitting the use of lenses of larger than normal aperture, thus doubling their efficiency (1917–21); two-color motion picture production on single-coated ordinary positive film (1914–18); two-color cine films made by cementing differently colored strips together, and bipack films (1917–19); trichromatic cine print in a single colloid layer (1916), and so on.

Frederic Ives was well into his seventies when he gathered his papers, *Scrap Books* and *Diaries* together and wrote, for his son and grandchildren, *The Autobiography of An Amateur Inventor*—the lively tale of a long and useful life. This he published privately in 1928. Then, characteristically, he went back to his laboratory and worked out the last of his inventions, the polychrome process of color photography, in which his genius provides an extremely simple method of making trichromatic prints from two negatives, which are produced by a single exposure on a Dupak film in an ordinary hand camera.

He was engaged with the commercial introduction of this process, and the establishing of a network of service stations for the processing of polychrome material, when his last illness came upon him. He died at Philadelphia, May 27, 1937, at the age of 81 years. Mr. Ives was twice married, and is survived by his son, Dr. Herbert E. Ives, a physicist of high repute in scientific research and invention.

# X-Rays in Agriculture

By Louis R. Maxwell and Sterling B. Hendricks
*Bureau of Chemistry and Soils,* U. S. Department of Agriculture*
*Washington, D. C.*

### *Introduction*

AGRICULTURE is one of the many activities in which use is made of radiography or the familiar production of shadows by variation in absorption of x-rays. The radiation as such can bring about biological changes in plants and animals and can be used for the production of new varieties. Atomic arrangements and intimately connected properties of soil constituents and products of the soil can be studied by diffraction of x-rays. This article attempts to describe some of the many applications according to these three methods. Since each has a technology of its own considerable space must be devoted to descriptions of products.

### *Radiography*

Inspection is one of the important operations in the preparation of food products for the market. In most cases there is the requirement that the product be uniform in quality for the value of a given shipment is greatly reduced by the presence of a few defective units. Citrus fruit growers in California[1] have been concerned with fruit damaged by cold weather; many oranges while normal in external appearance are pithy and juiceless as a result of freezing. A method therefore was needed for examining the inside of the fruit without injuring it and x-rays proved to be ideally adapted for this purpose. Good fruit, as shown in Fig. 1, has a dark fluoroscopic image while damaged fruit containing less juice has a lighter image. A typical inspection unit is shown in Fig. 2. Fruit is conveyed into the x-ray beam producing a shadow that is examined visually, poor fruit being manually ejected by the operator. The x-ray tube, which operates at 80 to 100 kv, is housed in a shock-proof compartment and the operators are protected from the radiation by lead glass observation windows. Depreciation of the apparatus and operating labor are the chief elements of cost, power consumption being at the rate of approximately one kilowatt. Four operators can inspect about two carloads of fruit in an eight hour day with the above unit.

Any lot of fruit that is outside the tolerance for the Sunkist or for the orchard run grade can be improved by removing the poor oranges detected on the x-ray inspection line. Shippers have been able to save large quantities of good fruit that could not have been salvaged by other methods. As an example, Whittier Citrus, a member of the California Fruit Growers' Exchange, examined one lot of navel oranges that had not been considered worth picking; thirty percent were removed and the remainder passed the Sunkist requirements. The grower thus at an ultimate economy to himself places a more desirable product on the market.

Strange objects sometimes find their way into packaged food products and these often result in expensive legal suits. Most foreign particles

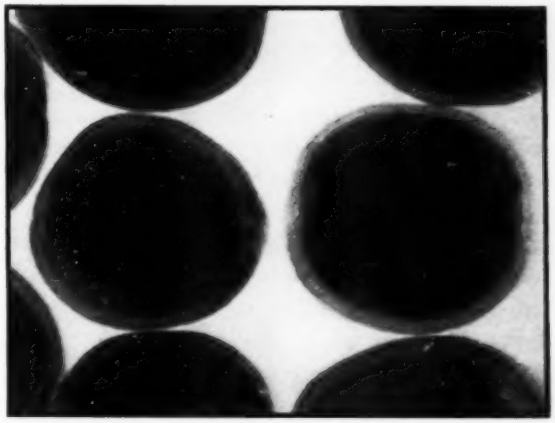


Fig. 1. Radiograph showing the mottled pattern of frost damaged oranges compared with normal fruit (Courtesy General Electric X-Ray Cooperation).

* Contribution from the Fertilizer Research Division.

are much more opaque to x-rays than are the products themselves and thus can readily be detected by a fluoroscopic method. Some canned

FIG. 2. X-ray fluoroscopic inspection unit for citrus fruit (Courtesy General Electric X-Ray Cooperation).

foods are regularly inspected in this manner which also guards against the presence of pits and partial filling of cans. A typical example is the inspection of salted peanuts.[2] Fig. 3a shows how difficult it is to see foreign objects placed in a tray of peanuts while Fig. 3b is an x-ray photograph of the same tray. One company regularly inspects its final product in this manner and while they find foreign objects in only one out of every 26,400 bags of peanuts they consider the inspection worth the cost.

Foreign objects in chewing tobacco, incipient pearls in oysters, holes in Swiss cheese, larvae in potatoes, textures of wood products, textiles, and papers, morphology of flowers, circulation of plant juices and many other topics have been studied by radiographic methods.[3] X-rays are also used quite extensively in veterinary work, the United States Army alone maintaining a number of x-ray outfits for examination of horses for bone fractures, foreign bodies, abscessed teeth, and infected sinus cavities. Progress of rickets in experimental animals is sometimes studied in this manner and one milling company has made extensive use of the fluoroscope in following bone growth in chickens as determined by feeding.

### *Irradiation*

Penetrating radiation such as x-rays can be absorbed by living matter with the liberation of photoelectrons or Compton recoil electrons. These primary particles dissipate their energy by ionization of the matter through which they pass, losing about 32 electron volts on the average for each ion pair formed. Biological changes result from this degradation of energy and cell patterns can be altered in minute or even lethal ways. A familiar application is in the therapy of cancer where preferential destruction of abnormal cells can frequently be brought about.

Scientific aspects of these phenomena are of interest to the geneticists, who are concerned with the manner in which living matter reproduces itself. It is well to introduce here a few concepts developed by the geneticist. Characteristics of plants and animals are determined by genes which are assumed to be submicroscopic bodies in the germ cell. Genes are considered to be the units of heredity and they are associated in a particular order in chromosomes which are usually microscopic in

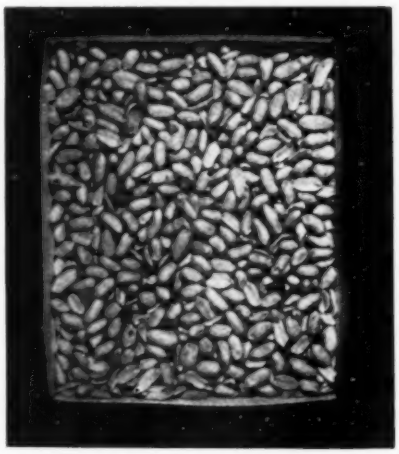


FIG. 3a

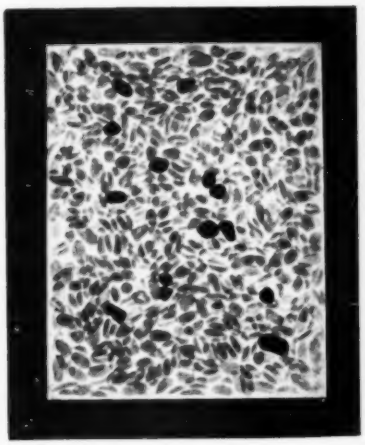


FIG. 3b

FIG. 3. a, Photograph of a tray of peanuts containing a few pebbles. b, Radiograph of the same tray. (Courtesy General Electric X-Ray Cooperation.)

size. Changes in offspring can be brought about by modification of the genes or by their rearrangement in the chromosomes. However, production of new varieties is usually a result of cross breeding which utilizes the latter factor. Morphological changes of the rapidly multiplying fruit fly *Drosophila melanogaster* have been studied most extensively. In this instance it has been shown that exposure to x-rays does increase the number of mutants or changed offspring. Moreover, some of the observed changes were shown to be the result of gene alteration.

Any artificial control of heredity is of agricultural interest particularly in the breeding of plants. For this reason many x-ray investigations[4] have dealt with this phase of agricultural problems. As an example, when the reproductive parts of tobacco, maize, and cereals in general are treated, very pronounced modifications are found. In one instance a variant of tobacco was selected for study and self-pollinated.[5] After six generations seven pure breeding types were secured. Some of these varieties were so changed from the original as almost to be new species; differing greatly from normal plants in habit of growth, leaf form, and colors of leaves and flowers. A typical dwarfed type of tobacco plant is shown in Fig. 4, this particular one being sterile. Dwarfed maize and vine-like barley are some of the typical forms resulting from x-radiation. Distribution of chlorophyll is frequently found to be modified in the mutant varieties, particularly for the grapefruit. The more altered plants often prove to be sterile.

FIG. 4. Dwarfed tobacco plant grown from seed treated with x-rays compared with a normal plant. (U. S. Bureau of Plant Industry, U. S. Bureau of Chemistry and Soils).

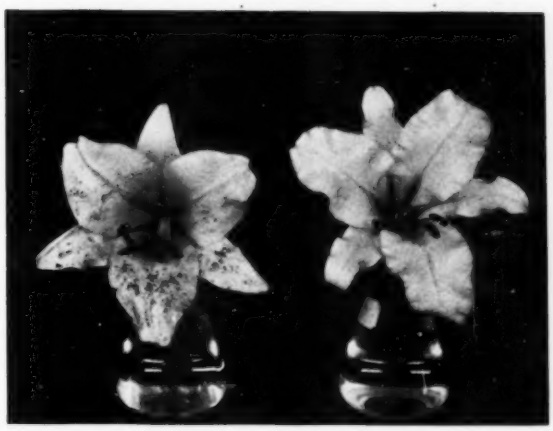


FIG. 5. An improved type of regal lily produced by x-ray treatment of bulbs (Courtesy General Electric Company). Left, normal; right, irradiated.

Since most of the changes produced by x-rays are destructive for some plant feature few of them are of practical importance. However, some decorative effect can be accomplished and for this reason flower culture is a possible field for application. A well-known success is that brought about by irradiation of the regal lily bulb (*Lilium Regale*).[6] Normally the florist removes the anthers from this beautiful flower since they break and shed their pollen on the white petals. The new variety does not have this objectionable feature as can be seen from Fig. 5. Bulbs producing this flower can be obtained by asexual reproduction of bulbs from the original plant. Similar experiments have been tried on tulips, petunias, narcissi (Fig. 6), and freesias, and changes have been produced in all cases.

In order to permit quantitative measurement of biological response to x-rays it is necessary to have some standard of dosage and this has

FIG. 6. Double paper-white narcissus grown from x-rayed bulb (courtesy Dr. C. P. Haskins).

been taken as the empirical unit called the "rontgen" or the "r" unit. The rontgen is defined by international agreement as the quantity of x-radiation that will produce in one cc of free air at standard conditions such a degree of conductivity that one e.s.u. of charge will be measured at saturation current. This is equivalent to an energy absorption of 0.107 erg per $cm^3$. A dosage of 500 to 1000 "r" units will produce a reddening of the skin or a so-called erythema dose. Some of the effects, such as the mutant of the regal lily described above, are observable for dosages as low as a few hundred "r" units which can be obtained in a fraction of a minute from many installations. Production of sterility in dormant germ cells such as rust spores often requires as much as 1,000,000 r units. Usually, effects are observed for dosages of a few thousand r units.

Instruments have been designed for measurement of the dosage and a typical apparatus, including the ionization chamber, is shown in Fig. 7. The specimen and the entrance diaphragm of the ionization chamber can be placed in equivalent positions with respect to the x-ray beam so that the correct dosage can be determined. A single layer of cotton gauze supports the specimen thus eliminating backward scattering. The ionization current is measured by means of a vacuum tube direct current amplifier circuit.

Biological response to x-rays is frequently interpreted according to the "hit theory,"[7] which assumes that the observed change is the result of the absorption of all or a portion of the energy of a high velocity electron traversing a sensitive volume within the organism. The theory can be formulated as follows; given $Z_0$ individuals each with one sensitive volume which requires $n$ successive hits to produce the desired effect then the number $Z$ surviving a dosage of $q$ rontgens is given by the following Poisson series:

$$Z/Z_0 = e^{-\alpha q}[1+\alpha q+(\alpha q)^2/2!+(\alpha q)^3/3!+\cdots +(\alpha q)^{n-1}/(n-1)!]$$

where $\alpha$ is the number of hits in the sensitive volume per rontgen.

Some single celled organisms have an exponential survival as a function of $\alpha q$, which indicates, according to the above equation, that a single hit in the sensitive volume brings about the observed effect. This is true for the killing

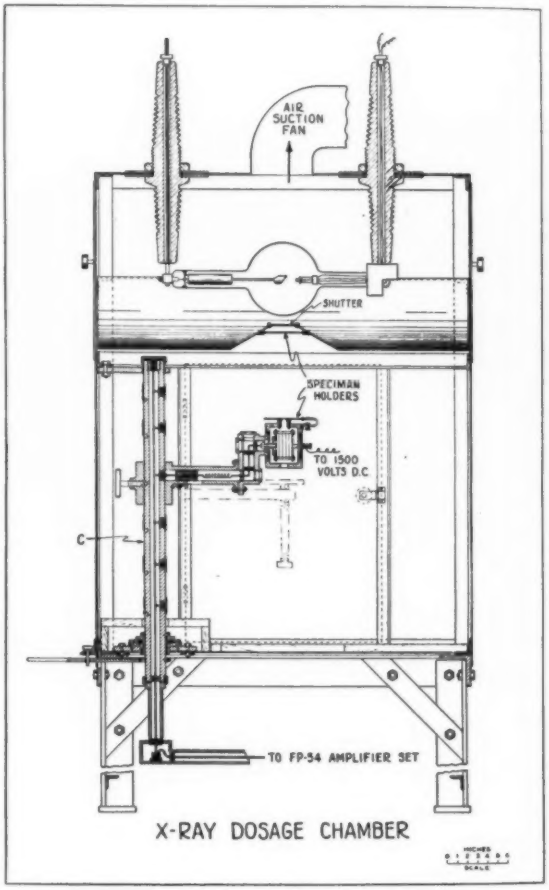



FIG. 7. A typical x-ray dosage chamber shown in elevation. (U. S. Bureau of Chemistry and Soils.)

of colon bacilli over the region between 0.5 and 4.0A[8] in wave-length. The calculated sensitive volume is about 1 percent of the entire cell volume which is cylindrical, being about $2\times10^{-4}$ cm long and $0.5\times10^{-4}$ cm in diameter.

Induced mutations such as the white eye of the fruit fly,[9] are linear functions of the dosage, independent of the wave-length. Lack of dependence upon wave-length can be interpreted as indicating that the mutations are produced by formation of ion pairs within the sensitive volume. These experiments are necessarily carried out under conditions where $\alpha q$ is small, which accounts for the observed linear relationships on the basis of a single hit. The sensitive volume for the white eye mutation computed for hits by single ion pairs comes out to be about $1.8\times10^{-18}$ cm$^3$ which corresponds to a spherical volume with a radius of about $7.6\times10^{-7}$ cm. These dimensions are close to the known sizes of protein molecules and of enzymes.

Multicellular organisms give S shaped survival curves[7] as a function of dosage and the delayed killing of corn seeds[10] illustrated in Fig. 8 is of

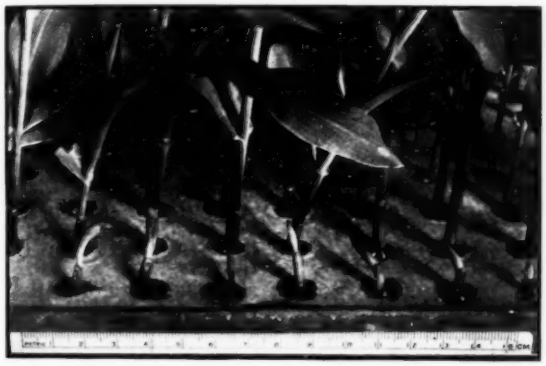


FIG. 8. Delayed killing of maize seedlings brought about by x-ray treatment of dry seeds (front row) compared with normal plants. (U. S. Bureau of Plant Industry, U. S. Bureau of Chemistry and Soils.)

this type. This class of curves can be interpreted according to the "hit theory" as requiring multiple hits in a single sensitive volume or hits in more than one sensitive volume. In the case of corn the first alternative is untenable since experiments[11] in which restricted areas of corn seeds were irradiated showed that there is no single sensitive volume of the dimensions required by the theory. S shaped curves, however, can also result from the inherent variability of biological materials.[12] Here it is merely necessary to consider the total energy absorbed by the specimen.

### *Diffraction*[13]

Study of crystalline solids and organic aggregates by x-ray diffraction is more closely restricted to physics and chemistry than are some of the examples mentioned above. Applications are chiefly of a research character rather than routine examination of a specific product. The objective usually is to aid in the understanding of a particular property that makes the material an economic one. Rational basis for modification rests upon such understanding.

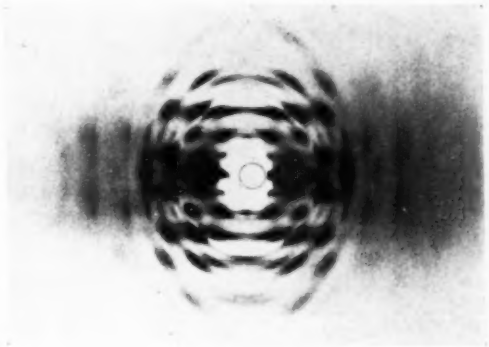


FIG. 9. X-ray diffraction pattern of ramie fiber (courtesy Dr. E. Sauter, Freiburg i. Br. Germany).

Structural parts of plants are composed chiefly of cellulose which not only serves the plant but is also widely used by man. The properties of cellulose that make it useful are its fibrous form which permits weaving into various shapes, the possible chemical modification of the cellulose, and the great tensile strength of the fibers. Chemical analysis shows that cellulose is made up of glucose molecules joined in indefinitely long groups; molecular weight determinations indicating that fifty or more molecules are joined in such groups. Microscopic examination suggests the presence of an additional material that fills in the gaps between groups. The high tensile strength probably results from the fact that large parts of the fiber are held together by

chemical bonds and thus are inherently strong within themselves. Other cohesive forces thus must only hold these units together and each unit has a large surface over which such forces can act.

X-ray diffraction shows how the glucose molecules are arranged in the fibers and how the

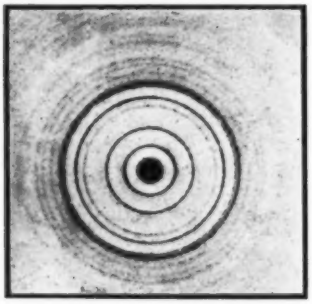


FIG. 10. Electron diffraction pattern of the clay mineral kaolinite. (U. S. Bureau of Chemistry and Soils.)

groups of molecules are related to the fiber axis. A diffraction pattern of ramie fiber is shown in Fig. 9. From this it can be seen that the fiber is highly organized and indeed closely approaches crystallinity. Closer analysis shows that all of the features of the pattern can be explained by an element of structure that is probably two glucose units in length along the fiber axis. The two glucose molecules are related by a twofold screw axis, a third molecule repeating the position of the first, a fourth that of the second .... It is this regularity that is associated with long chains rather than random aggregates. Fiber properties such as elasticity and swelling in solutions depend upon orientation of groups with respect to the fiber axis. Hard fibers like ramie and flax have a parallel orientation; cotton and some wood are spiral at a large angle. The orientation is more imperfect in many regenerated cellulose materials such as rayon.

Rubber owes its importance to its remarkable elastic extension and to its physical character after vulcanizing. Chemical analysis shows that it too is a condensation product but of isoprene molecules. Ordinary rubber shows no preferred orientation of individual units, but a well developed fiber diffraction pattern forms when it is stretched. It is this property of orientation that must be the basis of an adequate theory for the elastic properties of rubber.

Study of soil and its fertility is a major application of scientific methods to agriculture. The soil not only supports the plant and serves as a medium for bacterial action but it also supplies water and a portion of the plant food especially the inorganic part. Soil fertility depends in part upon base exchange capacity and ability to retain moisture. Study of soils in the laboratory has shown that these properties depend entirely upon the presence of degraded organic matter such as humus and finely divided inorganic compounds. It is in the study of the latter that x-ray diffraction finds application.[14] Sedimentation experiments show that a considerable portion of the soil is submicroscopic in dimensions and in fact approaches colloidal behavior. Diffraction patterns both x-ray and electron of these fractions show that they are not indefinite amorphous mixtures but rather contain only a few characteristic compounds. Further studies have shown that these compounds are predominantly clay minerals.

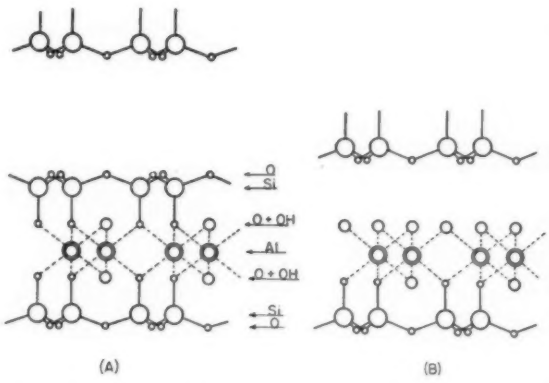



FIG. 11. Compositions of layers in the clay minerals, (A) montmorillonite; (B) kaolinite, halloysite, nacrite, and dickite; circle sizes as shown are not related to ionic radii.

Three distinct types of minerals are often present; those related to kaolinite, $Al_2O_3 \cdot 2SiO_2 \cdot 2H_2O$, to montmorillonite, $Al_2O_3 \cdot 4SiO_2 \cdot H_2O$, and to a mica-like mineral containing potassium.

An electron diffraction pattern of the clay fraction of the common Cecil soil is shown in Fig. 10. Study of these and of x-ray diffraction patterns lead to determinations of structures for

the minerals. They are found to be made up of layers of the type shown diagrammatically in Fig. 11 and by the space model shown on the front cover. Individual layers have strong internal Coulomb forces but only weak van der Waals forces operate between separate layers. As a result the minerals can be degraded to very small particle size and thus give large surfaces for exchange of basic ions. This possibility of exchange is further a result of the structure in which $Al^{+++}$ ions can be replaced by $Mg^{++}$, $Ca^{++}$, $Na^{+}$, etc. In montmorillonite oxygen ions only form the upper and lower surfaces of the layers while in kaolinite minerals hydroxyl ions of one surface are adjacent to oxygen ions of another. Study of many minerals has shown that the attractive forces in the second case are greater than those of the first. Montmorillonites, which form the so-called bentonites, have the unusual property of swelling when placed in water. This water goes between the layers as a result of the small attractive forces and in so doing it increases the stability of the lattice.

X-ray diffraction also finds application in fertilizer technology, particularly that of the compounds of phosphorus that are essential plant foods. Phosphate rock, the principal source of such compounds, must be changed chemically to be available as plant food. The most economic manner of accomplishing this is the subject of intensive study at the present time and various methods are being tried. Compounds present in intermediate and final products often can be identified only from their diffraction patterns. This is particularly true of those materials resulting from the action of steam on the rock at high temperatures. Here it is found that any compound produced will be available as a plant food as long as it does not have the lattice of the mineral apatite. This mineral is the predominant constituent of the rock itself and is also a bone material.

### *Remarks*

These few examples will perhaps suffice to show the possibilities of applying x-rays to the problems of agriculture. Physicists have contributed to the basic information necessary for these developments. They have also supplied the practical apparatus at moderate cost and they have made it available for common use. It can safely be predicted that there will be increased application along the general lines outlined above.

We are indebted to Mr. J. R. MacRill of the California Fruit Growers Exchange and to the General Electric X-Ray Cooperation for information on inspection of oranges. Dr. C. P. Haskins and the General Electric Company kindly furnished us with much material. The Bureau of Plant Industry of the U. S. Department of Agriculture cooperated on some of the biological work mentioned.

### *Bibliography*

1. *The California Citrograph* (1937), p. 142, 291, 495.
2. Arachis H. Leguminosa, Confectioners Journal, June (1937).
3. "Symposium on Radiography and X-ray Diffraction Methods," Am. Soc. for Testing Materials 92 (1936).
4. See various sections in *Biological Effects of Radiation*, edited by B. M. Duggar (New York, 1936).
5. U. S. Dept. of Agriculture, *Yearbook of Agriculture*, (1937), p. 913.
6. Chester N. Moore, U. S. Plant Patent No. 165.
7. See for instance, R. Glocker, Zeits. f. Physik **77**, 653 (1932).
8. R. W. G. Wyckoff, J. Exp. Med. **52**, 769 (1930). For further interpretation see R. Glocher, reference 7; also F. L. Mohler and Lauriston S. Taylor, Nat. Bur. Stand. J. Research **13**, 677 (1934).
9. C. P. Haskins, Proc. Nat. Acad. Sc. **21**, 561 (1935). For a review article on the experimental production of mutations see N. W. Timoféëff-Ressovsky, Biol. Rev. and Biol., Proc. Camb. Phil. Soc. **9**, 411 (1934). See also J. W. Gowen, *Biological Effects of Radiation* (Duggar) Vol. 2 (New York, 1936), p. 1311.
10. G. N. Collins and L. R. Maxwell, Science **83**, 375 (1936).
11. L. R. Maxwell, Phys. Rev. **51**, 375 (1937).
12. George C. Whipple, J. Frank. Inst. **182**, 37 (1916); A. J. Clark, *The Mode of Action of Drugs on Cells* (London, 1933), pp. 97–130.
13. For a general discussion on structure of crystals, note any standard reference such as "Struktur-Bericht," Zeits. f. Krist.
14. S. B. Hendricks and W. H. Fry. Soil Science **29**, 457 (1930).

# Order and Disorder in Alloys

By H. A. BETHE

*Cornell University*

*Ithaca, New York*

## I

THE increasing importance of alloys in industry has done much to stimulate research about their structure. Foremost among the methods used in this research is, of course, x-ray analysis, but also electrical conductivity, thermal properties, etc., give valuable information.

It has long been known that there are two main classes of alloys, homogeneous (one phase) alloys and eutectics (two phase systems). In the homogeneous alloys, the concentration of each of the metals contained in the alloy is constant throughout the alloy. The inhomogeneous alloys consist of small crystals some of which contain more of one component, while others contain more of the other. Such eutectics are formed when the mutual solubility of the components of the alloy is limited and the concentration in the given alloy is beyond the solubility limits.

We shall be primarily concerned with the homogeneous alloys. Here again two possibilities arise: The atoms of different kinds may be distributed at random over the lattice points of the crystal, or they may be in an ordered distribution, forming a "superlattice." A superlattice in the strict sense is only possible for a rational composition of the alloy, the simplest case being when there are equally many atoms of each of two kinds, $A$ and $B$. Let us consider an alloy of this composition which crystallizes in a cubic body-centered lattice. Then perfect order is obtained if all the corners of the lattice cubes are occupied by $A$ atoms, and all the centers by $B$ atoms; then the $A$ atoms form a simple cubic lattice among themselves, and so do the $B$ atoms among themselves. On the other hand, in a random distribution, corners as well as centers would be occupied by $A$ atoms as often as by $B$ atoms.

If the composition is not exactly "rational" but very closely such, we can still have a superlattice. If we have, e.g., an alloy containing 49 percent $B$ atoms and 51 percent $A$ atoms, again crystallizing in a cubic body-centered lattice, we may fill all "corners" with $A$ atoms, and put the $B$ atoms as well as the remaining (1 percent) $A$ atoms into the centers of the cubes.

Besides the simplest superlattice which is obtained for a 1 : 1 composition of the alloy, there are other, slightly more complicated superlattices for other compositions such as 1 : 2, 1 : 3, 1 : 5, 1 : 7, 3 : 5, etc. For a given crystal structure, not all of these are possible. The frequently occurring cubic body-centered structure permits simple superlattices of cubic symmetry for the compositions 1 : 1 (described above), 1 : 3, 1 : 7, 3 : 5. If we have 3 $A$ atoms to each $B$ atom (3 : 1 composition), we may fill all the cube centers with $A$ atoms and make the corners alternately $A$ and $B$ so that the $B$ atoms form a cubic face-centered lattice among themselves. With 7 $A$ atoms to one $B$, the $B$'s may be arranged in a cubic body-centered superlattice of twice the lattice constant of the main lattice. The compositions 1 : 2 and 1 : 5, on the other hand, do not lead to any simple superlattice.

If the main lattice is simple cubic, again the compositions 1 : 1, 1 : 3, 1 : 7 give simple superlattices, the $B$ atoms forming respectively a face-centered, body-centered and simple cubic superlattice having twice the spacing of the main lattice.

For the cubic face-centered structure, only the compositions 1 : 3 and 1 : 7 lead to simple superlattices ($B$ atoms simple cubic and face-centered, respectively). With the composition 1 : 1 it is impossible to construct a superlattice of cubic symmetry. The only possibility is to arrange the $A$ and $B$ atom in alternating planes; by this arrangement one axis (perpendicular to the planes) will be distinguished from the others so that the structure will become tetragonal rather than cubic. The lattice spacing will, then, also be different along the different axes, in general smaller along the tetragonal axis. This is exactly what happens in the alloy Cu Au (spacing 3.70 angstrom along tetragonal axis, 3.94 angstrom in planes). A 1 : 2 composition permits the formation of a hexagonal superlattice, with the $A$ and $B$ atoms in planes perpendicular to the cubic diagonal, arranged in the order $AABAAB\cdots$.

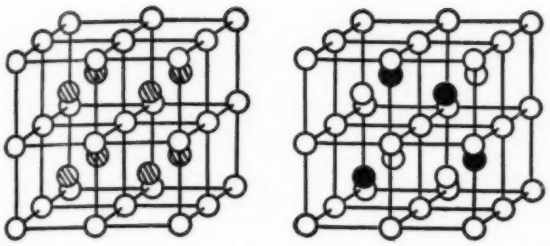


FIG. 1. Superlattices of composition 1 : 1 (left) and 1 : 3 (right) in a cubic body-centered crystal. Open circles, *A* atoms; filled circles, *B* atoms.

## II

In order that a superlattice be formed it is, of course, necessary that this ordered arrangement of atoms be energetically more favorable than the random distribution. Quite generally, in a superlattice, unlike atoms are close together and atoms of the same kind more distant from each other. Therefore it is necessary that the attraction between unlike atoms be greater than the (average) attraction between like atoms. If the opposite is the case, there will be a tendency for the various kinds of atoms to be segregated from each other and we shall find a eutectic mixture rather than a superlattice.

Even if its lower energy favors the superlattice, it will not be stable at all temperatures. For there is only one distribution of the atoms among the lattice points which corresponds to a given superlattice while there is an enormous number of possible disordered distributions. At high temperatures, the disordered state will become more stable while at low temperatures the state of lowest energy will be the stable one, i.e., the superlattice. We thus expect a transition temperature at which the superlattice ceases to be stable; we shall call it the critical temperature, $\Theta$.

The critical temperature will be a measure of the force which tends to order the crystal; the greater this force, the higher $\Theta$. In general, the ordering force will be the greater the simpler the composition of the alloy; i.e., the characteristic temperature will be highest for a 1 : 1 superstructure, lower for the 3 : 1 case and still lower for 7 : 1 or 5 : 3. The $\Theta$ for 1 : 1 superstructures will often be higher than the melting point of the substance; in that case the superlattice will persist up to the melting point. In many other cases, the critical temperature will be below the melting point so that an actual transformation from the disordered to the ordered state can be observed when the alloy is cooled.

The formation of the superlattice necessitates a complete rearrangement of the atoms in the alloy. Such a rearrangement usually requires a considerable time. Thus, if the alloy is cooled very rapidly (quenched), there will in general not be enough time for the establishing of the superlattice; then the disordered state (solid solution) will persist at temperatures at which it is thermodynamically unstable. Since the rearrangement of atoms is very much slower at low temperatures than at high ones, an alloy once quenched will remain in its disordered state almost indefinitely as long as it is kept at sufficiently low temperatures. In order to obtain the ordered modification, slow cooling (annealing) is essential; the alloy must be kept for a long time at a temperature somewhat below the transformation temperature $\Theta$. The time required for the annealing varies with $\Theta$; for high $\Theta$ (about 1000°C, with a very simple composition of the alloy) it may be a small fraction of a second so that the quenching rather than the annealing becomes the difficult process; for low $\Theta$ (room temperature) it may become of the order of many years or even millions of years so that there is no hope of ever forming superlattices of complicated composition* (such as 1 : 7) and correspondingly low transformation temperature. Only for a rather restricted class of alloys, of medium $\Theta$ ($\sim$300–700°C), the annealing time is in the accessible range, say, from a second to a year, and only for those it is possible to obtain both the ordered and the disordered modification at room temperature, by either annealing or quenching.

## III

Considerable effort has been spent in obtaining a theoretical understanding of the mechanism of the transformation from the ordered to the disordered state. Bragg and Williams[1] were the first to fully recognize that this transformation will ordinarily be gradual rather than sudden. Taking again, for definiteness, the case of a cubic

* This does not affect the superlattices of the Hume Rothery type such as $\gamma$-brass (5 atoms of Cu to 8 atoms of Zn) whose energy is particularly low according to the electron theory of metals, cf. Jones, Proc. Roy. Soc. **A153**, 622 (1936).

body-centered crystal containing equally many $A$ and $B$ atoms, we obtain perfect order when all corners are occupied by $A$ atoms, all centers by $B$, while perfect disorder means that 50 percent of the corners are $A$, 50 percent $B$. But between these extreme cases, any intermediate state is possible, and the effect of raising the temperature will be to decrease gradually the number of $A$ atoms in corners from 100 percent to 50 percent. These considerations lead immediately to the definition of the degree of (long range) order:

$s$ = The number of $A$ atoms in corners minus the number of $B$ atoms in corners, divided by the total number of corners (1)

$s=1$ and $-1$ correspond to a perfect superlattice, $S=0$ to disorder.

The fundamental problem in the theory of superlattices is then the calculation of the degree of order as a function of the temperature $T$ (always measured in degrees Kelvin). For this purpose, we must first know the forces acting between the atoms. Quantum theory shows quite generally that the forces between atoms decrease rapidly as the distance separating them increases. Thus it will be a good approximation to assume forces only between nearest neighbors. These forces are conveniently expressed by giving the interaction energy (potential energy) of two atoms at the actual distance of nearest neighbors in the crystal. Let the interaction energy of two unlike neighbors be $V_{AB}$, that of two like neighbors $V_{AA}$ or $V_{BB}$, respectively. Then the necessary condition for the formation of a superlattice is

$$V_{AB}<\tfrac{1}{2}(V_{AA}+V_{BB}). \tag{2}$$

In order to find the energy of a given distribution of $A$ and $B$ atoms, we must know how many pairs of unlike neighbors ($AB$) and how many pairs of like neighbors ($AA$ or $BB$) occur in the crystal. Let $N$ be the total number of $A$ atoms (equal to the number of $B$ atoms) and $z$ the number of nearest neighbors of each atom ($z=6$ for simple cubic, 8 for body-centered, 12 for face-centered lattice). Then, if $M$ is the number of pairs of unlike neighbors $AB$, there will be $\frac{1}{2}(Nz-M)$ pairs of $AA$ neighbors and equally many $BB$ neighbors. Then the energy of the whole crystal is

$$MV_{AB}+\tfrac{1}{2}V_{AA}(Nz-M)+\tfrac{1}{2}V_{BB}(Nz-M) = \text{const}-VM, \tag{3}$$

where $$V=\tfrac{1}{2}(V_{AA}+V_{BB})-V_{AB}>0. \tag{3a}$$

For perfect order all neighbors of $A$ atoms are of kind $B$ and *vice versa* so that $M=Nz$. For complete disorder, the neighbors of $A$ atoms will be $A$ as often as $B$ atoms so that $M=\frac{1}{2}Nz$. We define the "local order"

$$\sigma=2M/Nz-1. \tag{4}$$

The energy of the crystal is then

$$U=\text{const.}-\tfrac{1}{2}NzV\sigma. \tag{5}$$

The local order $\sigma$ is not uniquely determined by the long range order $S$. It is true that, for perfect long range order, the local order must also be perfect, but it is very well possible to have no long range order at all but still have some local order. Local order means simply that the neighbor of an $A$ atom is more likely to be a $B$ atom than another $A$, and this is quite reconcilable with an equal distribution of the $A$ atoms among "corners" and "centers" of cubes. We may even have zero long range order and almost

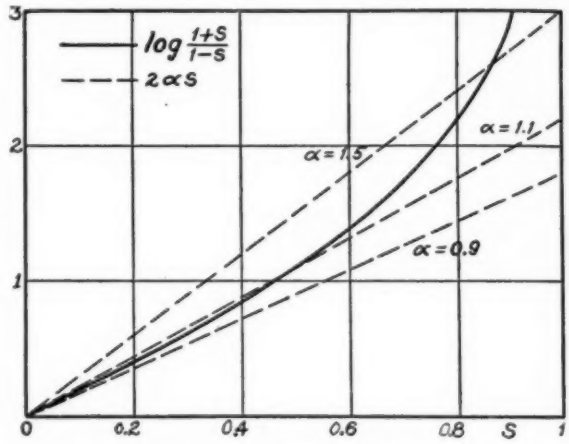



FIG. 2. Graph for the determination of the long range order $s$ as a function of the temperature, approximation of Bragg and Williams. The curve gives $\log (1+s)/(1-s)$, the straight lines give $2\alpha s$ for $\alpha=0.9$, 1.1 and $\alpha=1.5$. $\alpha=0.9$ corresponds to a temperature above the critical temperature, there is no intersection and therefore zero long range order. For $\alpha=1.1$ (temperature 10 percent below the critical temperature) an intersection is found at $s=0.5$ representing the long range order at that temperature. For $\alpha=1.5$ (temperature considerably below $\Theta$), the intersection is at almost complete order.

complete local order. We need only divide our crystal into two equal parts each of which is perfectly ordered while the two parts are "out of phase" with each other. Thus there is no unique correspondence between long range order and local order. However, there is, of course, some correlation between the two, *viz.* distributions of high long range order will, on the average, also show greater local order. The energy is uniquely determined by the local order and is lowest for high local order; since the local order increases, on the average, with the long range order, the distributions of great long range order will in general have low energy.

The concepts of local order and long range order have an obvious application to the theory of solids and liquids. In a liquid, there is very definite local order, but no long range order; in a solid, we have long range as well as local order. The positions of two atoms in a crystal, however distant from each other, are exactly correlated; if the position of one atom and the direction of the crystal axes are given, we can predict the position of another atom at a distance of thousands of atomic diameters with an accuracy of a small fraction of an atomic diameter; and this remains true even if lattice vibrations are considered. In a liquid, there is no correlation whatever between distant molecules, but neighboring molecules show a very definite order, with respect to their distances as well as with respect to the number of molecules surrounding a given one. This local order may be still noticeable among second and further neighbors of a given molecule and it gives rise to the well-known rings in x-ray diffraction. But, however good the local order, there will never be any long range order in a liquid. The presence or absence of long range order seems to be the most characteristic difference between a solid and a liquid.

## IV

The first theory of the stability of superlattices, as a function of temperature, was given by Bragg and Williams[1] and extended by Williams.[2] A more refined method was proposed by Bethe[3] and applied by Peierls[4] and Chang[5] to various cases. Neither of these methods is quite satisfactory because the assumptions made are not quite clear. This state of affairs has been remedied in a recent paper by Kirkwood[6] which is based entirely on standard concepts of statistics. Kirkwood has shown that Bragg and Williams' results are obtained as a first approximation from his method, and also Bethe's method can be justified on the basis of Kirkwood's calculations. We shall follow Kirkwood in our discussions.

Consider all the distributions of long range order $s$. If we refer again to a cubic body-centered crystal, there will be (cf. Eq. (1))

$$R=\tfrac{1}{2}N(1+s) \tag{6}$$

atoms of kind $A$ and

$$W=\tfrac{1}{2}N(1-s) \tag{6a}$$

$B$ atoms in the corners of cubes while the remaining $W$ $A$ atoms and $R$ $B$ atoms are in the centers. There will be

$$N!/R!\,W!$$

different ways of distributing the $A$ and $B$ atoms among the $N$ corner positions, and equally many distributions of the atoms among the $N$ center positions. The total number of distributions is therefore

$$n=(N!/R!\,W!)^2. \tag{7}$$

The probability of finding a given distribution of the atoms among the lattice ponits depends on the energy of the crystal through the Boltzmann factor

$$e^{-U/kT} \tag{8}$$

where $k$ is Boltzmann's constant and $T$ the absolute temperature. With $U$ given by (5), the probability is proportional to

$$p=e^{N\alpha s} \tag{9}$$

with

$$\alpha=zV/2kT. \tag{10}$$

If all the distributions of given long range order $s$ had the *same local order* $\sigma(s)$, the probability of finding a distribution of long range order $s$ would be (cf. (6), (8)) proportional to

$$P=(N!/R!\,W!)^2\,e^{N\alpha s}. \tag{11}$$

At any given temperature, i.e., for any given $\alpha$, $P$ will have a maximum for a certain value of the long range order $s$ which is the most probable long range order for that temperature. The condition that (11) be a maximum, together with the definitions (6) (6a) of $R$ and $W$, lead to the

equation

$$\left(\frac{R}{W}\right)^2 = \exp\left(N\alpha\frac{d\sigma(s)}{dR}\right) = \exp\left(2\alpha\frac{d\sigma(s)}{ds}\right), \quad (12)$$

$$\log\frac{1+s}{1-s} = \alpha\frac{d\sigma(s)}{ds}. \quad (13)$$

The solution of this equation gives the order $s$ as a function of $\alpha$, and therefore, through (10), of the temperature.

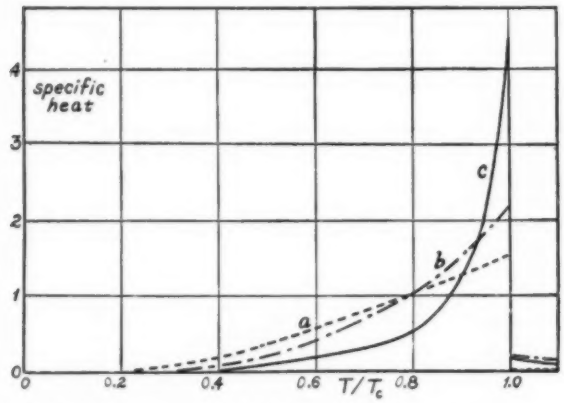



FIG. 3. The specific heat as a function of temperature, for simple cubic structure and 1 : 1 composition. Curve $a$, approximation of Bragg and Williams; curves $b$ and $c$, more exact theory. Curve $b$ is calculated without interaction between second nearest neighbors, $c$ with such an interaction, equal to 0.094 of the interaction between nearest neighbors.

As a first approximation, we may insert for $\sigma(s)$ the *average* local order for a given long range order, averaged over all the $n$ (cf. (7)) distributions of order $s$. This average local order is easy to calculate. According to (4), the local order is given if we know the number, $\frac{1}{2}(Nz-M)$, of pairs of like neighbors $AA$ in the crystal. The number of $A$ atoms in "corners" is $R$, the fraction of cube centers occupied by $A$ atoms is $W/N$. The number of neighbors of each corner $A$ atom is $z$, and in the average over all corner $A$ atoms and over all distributions of long range order $S$, there will be $zW/N$ $A$ atoms among these neighbors. Therefore the number of $AA$ pairs is

$$\tfrac{1}{2}(Nz-M) = zRW/N \quad (14)$$

and, according to (4), (6), (6a)

$$\sigma(s) = 1-4(RW/N^2) = s^2. \quad (15)$$

Inserting this into (13), we find

$$\log(1+s)/(1-s) = 2\alpha s. \quad (16)$$

This equation was first derived by Bragg and Williams using a somewhat different argument. It is most easily solved graphically. Fig. 2 gives the left-hand side as a function of $s$; the curve starts out at the origin with slope 2 and becomes then increasingly steeper. There will be an intersection with the straight line $2\alpha s$ if $\alpha>1$. The intersection will be at small values of $s$ if $\alpha$ is slightly greater than one, and will move towards $s=1$ for greater $\alpha$. According to the definition (10) of $\alpha$, this means that *the order will be nearly perfect* ($s\approx 1$) *for low temperatures; it will gradually decrease with increasing temperature and will reach zero at the critical temperature*

$$\Theta = \tfrac{1}{2}zV/k, \quad (17)$$

which corresponds to $\alpha=1$. Above the critical temperature, the long range order will remain zero.

By expanding the left-hand side of (16) in powers of $s$, it follows that, for $\alpha$ slightly greater than one,

$$s = (3(\alpha-1))^{\frac{1}{2}} \approx (3(\Theta-T)/\Theta)^{\frac{1}{2}}, \quad (18)$$

i.e., the curve representing the long range order as a function of the temperature approaches zero with vertical tangent at the critical temperature (cf. Fig. 5). Since in our theory the short range order is $\sigma(s)=s^2$, the energy of the crystal is, according to (5), (17)

$$U = \text{const.} - 3Nk(\Theta-T) \quad (19)$$

for temperatures $T$ slightly below the critical temperature $\Theta$. This energy increases rapidly with the temperature, which gives rise to an extra specific heat of the crystal; this is, per atom

$$c = \frac{1}{2N}\frac{dU}{dT} = \frac{3}{2}k \quad (20)$$

since there are $2N$ atoms in the crystal. This extra specific heat (3 cal./mole), which is added to the ordinary specific heat of about 6 cal./mole, has actually been measured. Above the critical temperature the specific heat is zero, in our approximation (see, however, below). Thus there is a *jump of the specific heat at the critical temperature*

but no latent heat is developed (cf. Fig. 4). A transformation of this type is commonly called a phase change of the second kind.

## V

The Bragg and Williams approximation which was outlined in the preceding section is based upon the assumption that the local order is a unique function of the long range order. As we have seen in III, this is actually not the case. Owing to the Boltzmann factor, $e^{N\alpha\sigma}$, the distributions of higher local order will be relatively more probable. Therefore the most probable local order will, for a given long range order, actually be greater than $s^2$ which was the average value calculated in IV (Eq. (15)). Furthermore, it will be a function of the temperature as well as of $s$, so that we must write $\sigma=\sigma(s, \alpha)$. The difference $\sigma(s, \alpha)-s^2$ will be the greater the lower the temperature because of the greater importance of the Boltzmann factor, and it will in general be the smaller the greater $s^2$ is itself. This latter fact has the consequence that the long range order will in general be smaller, at a given temperature, than the value given by the Bragg and Williams theory.

The main differences between the results of the more exact theory and the Bragg and Williams approximation are:

(1) The critical temperature is lower. For a simple cubic lattice ($z=6$), the critical temperature is

$\Theta=3\ V/k$ (Bragg and Williams),
$\Theta=2.37\ V/k$ (more exact theory).

(2) The setting in of long range order below the critical temperature is sharper. For the simple cubic lattice,

$s=1.73\ ((\Theta-T)/\Theta)^{\frac{1}{2}}$ Bragg and Williams,
$s=2.32\ ((\Theta-T)/\Theta)^{\frac{1}{2}}$ more exact theory.

(3) The extra specific heat below the critical temperature is larger, and depends on the number of neighbors $z$. We have for the extra specific heat per atom

$c=1.50\ k$ for *any* crystal lattice, according to Bragg and Williams,
$c=2.14\ k$ for the simple cubic lattice in the more exact theory.
$c=1.95\ k$ for the body-centered lattice in the more exact theory.

(4) There is some extra specific heat above the critical temperature because there is some local order left even for $s=0$, and this local order decreases with a further increase of the temperature. For the simple cubic lattice, the extra specific heat above the critical temperature is about one-tenth of that below.

(5) The total heat evolution due to the disordering process, from zero to infinite temperature, is equal to $\frac{1}{4}zV$ per atom (cf. (5)). According to the Bragg and Williams approximation, all this heat is evolved below the critical temperature. Actually, there is a residual local order at $T=\Theta$ which is slightly greater than $1/z$ (for $z=6$, it is 24 percent). Therefore a part of the heat evolution (somewhat more than $\frac{1}{4}V$) takes place above the critical temperature, somewhat less than $\frac{1}{4}(z-1)V$ per atom is evolved below the critical temperature.

(6) Since the total heat evolution below the critical temperature is less than in the Bragg-Williams approximation, and the specific heat just below $\Theta$ is greater, the extra specific heat extends over a smaller range of temperature. In the Bragg-Williams approximation, the specific

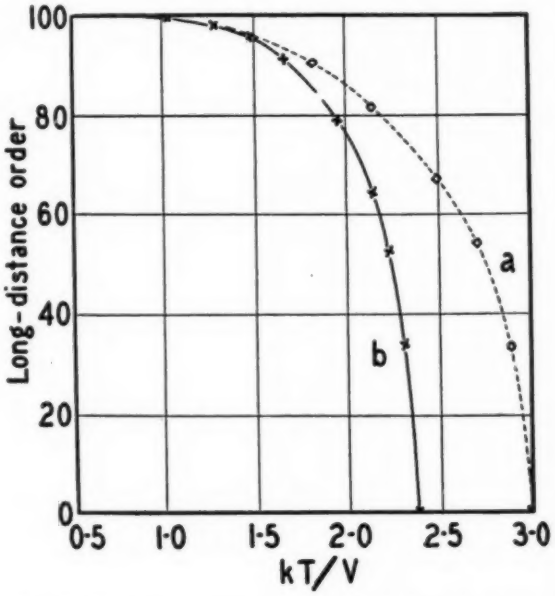



FIG. 4. The long range order as a function of the temperature (simple cubic, 1 : 1 composition). *a*, Bragg and Williams; *b*, more exact theory.

heat at 0.7 $\Theta$ is about one-half of the specific heat at the critical temperature itself. In the more accurate theory, the specific heat falls to one-half already at about 0.82 $\Theta$.

Figure 3 gives a comparison of the specific heats in the Bragg-Williams approximation and in the more accurate theory, for the simple cubic lattice. Fig. 4 gives the long range order in the two approximations, Fig. 5 the local order. It is evident that the transformation is sharper in the more accurate theory.

## VI

Chang[5] has investigated the influence of the interaction between second nearest neighbors on the order in alloys. Assuming that this interaction has the same sign as that between nearest neighbors, the second neighbor interaction will oppose the formation of a superlattice. For in the superlattice second neighbors will be of the same kind and this means higher interaction energy. Thus the critical temperature will be lower than for interaction between nearest neighbors only. With a second neighbor interaction $V_2$ equal to 0.094 times the interaction $V$ between nearest neighbors, the critical temperature is 1.8 $V/k$ as against 2.4 $V/k$ when the second neighbors are neglected. For this case, $V_2=0.094\ V$, the extra specific heat is given as a function of $T/\Theta$ in Fig. 3, curve *c*. The curve is seen to be much steeper than without the interaction between second neighbors, rising to 4.14 $k$ per atom just below the critical temperature (2.14 with $V_2=0$). The whole transformation occurs in a much narrower temperature region, the specific heat falling to one-half its maximum at $T=0.95\Theta$ ($0.82\Theta$ for $V_2=0$). This can be understood qualitatively.

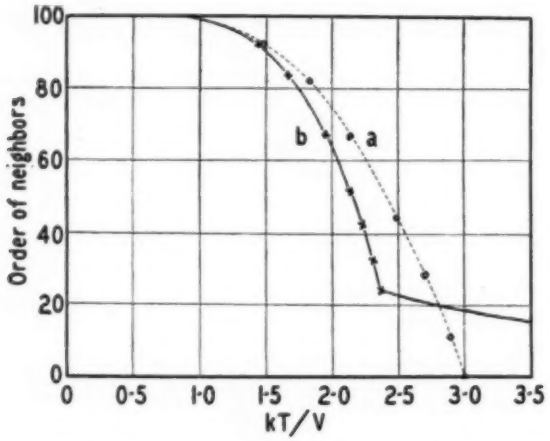



FIG. 5. The short range order as a function of temperature
*a*, Bragg and Williams; *b*, more exact theory.

With increasing interaction between second neighbors, the specific heat at the critical temperature becomes larger, and it is very sensitive to the exact value assumed for the second neighbor interaction.* Experimentally,[7] the specific heats are rather large ($\approx 7k$ per atom) and the transformation correspondingly sharp, indicating a considerable interaction between second neighbors. For very large second neighbor interaction ($V_2>0.23\ V$ for simple cubic lattice), there will be a latent heat at the critical temperature, besides the extra specific heat below $\Theta$. For still larger $V_2$, the formation of a superlattice becomes impossible.

Peierls[4] has investigated the order as a function of temperature for the superlattice of composition 1 : 3 in a face-centered cubic lattice. $Cu_3Au$ is an example of such a lattice (corners of cubes occupied by Au atoms, face centers by Cu). Peierls finds that the order sets in very suddenly; the long range order jumps from about 90 percent to zero at the critical temperature, and 80 percent of the heat evolution below $\Theta$ takes place at the critical temperature itself, as a latent heat. The reason for this very sharp transition is probably that the interaction between the atoms at the face centers opposes the formation of a superlattice, similarly to the effect of second neighbors in the ordinary (1 : 1) superlattice discussed above.

## VII

The formation of a superlattice has an influence on almost every physical property of an alloy. Thus far, we have only discussed the thermal properties (latent heat, extra specific heat) because they give the most direct information on the local order as a function of temperature. The most direct evidence for the long range order is, of course, obtained from x-ray analysis. If a superlattice is formed, there will be additional x-ray reflections between the ordinary ones. The

* The interaction between third neighbors will tend to decrease the maximum specific heat again and make the transformation less sharp, etc.

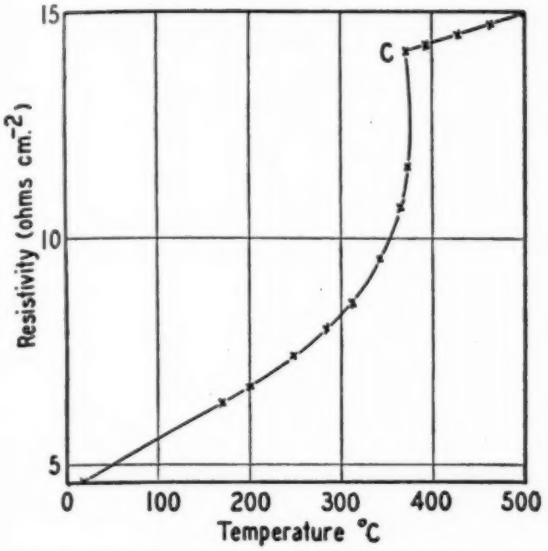



FIG. 6. The electrical resistivity of CuZn as a function of temperature, according to measurements of Sykes and Evans.

intensity of these reflections will be proportional to the square of the long range order.

The electric conductivity shows a very striking dependence on the order in the lattice, especially at low temperatures.[8] According to the quantum theory of conductivity,[9] electrons in a perfectly periodic lattice move without resistance. The actual resistance of a metal is due (1) to the thermal vibrations of the atoms which give deviations from periodicity, and (2) to disorder. Therefore the ordered alloys will have much lower resistance than the disordered ones, and there will be a sharp increase of the resistance when the order disappears at the critical temperature (Fig. 6).

As already mentioned in II, a certain time is required for the formation of a superlattice. When an alloy is quenched, order will be established only partially or, sometimes, not at all. If the order obtained by quenching is measured (e.g. by means of the resistivity) as a function of the cooling rate, this will give an idea about the time required for ordering. It is to be expected that this time will be very much longer for the establishing of perfect order than for partial order.

## VIII

The formation of superlattices is the simplest example of a "cooperative phenomenon," i.e. a process which depends upon the interplay of a large number of atoms rather than on the individual atoms. Other well-known cooperative phenomena are melting and ferromagnetism. The theory of superlattices is simpler than that of either of the other phenomena. In contrast to ferromagnetism which cannot be understood without a thorough quantum mechanical investigation, the theory of superlattices is purely a problem of classical statistics. The main obstacle to a rigorous theory of melting is the difficulty of an adequate geometrical description of the liquid state while the geometrical description of superlattices is extremely simple. Because of its comparative simplicity, the problem of order and disorder in alloys has been more completely solved than the other cooperative phenomena, although much remains to be done especially as regards quantitative results. The solution of the order-disorder problem throws, no doubt, also light on the other cooperative phenomena; in fact, Bitter[10] has shown that an adequate phenomenological description of ferromagnetism may be given on the lines of the superlattice theory.

### *Bibliography*

1. Bragg and Williams, Proc. Roy. Soc. **A145**, 699 (1934); **A151**, 540 (1935).
2. Williams, Proc. Roy. Soc. **A152**, 231 (1935).
3. Bethe, Proc. Roy. Soc. **A150**, 552 (1935).
4. Peierls, Proc. Roy. Soc. **A154**, 207 (1936).
5. Chang, Proc. Roy. Soc. **A161**, 546 (1937).
6. Kirkwood, J. Chem. Phys. **6**, 70 (1938).
7. Sykes, Proc. Roy. Soc. **A148**, 422 (1935).
8. Sykes and Evans, J. Inst. Metals **58**, 255 (1936).
9. Mott and Jones, *Theory of Metals*, (Clarendon Press, Oxford, 1935). Seitz and Johnson, J. App. Phys. **8**, 186, 246 (1937).
10. F. Bitter, *Introduction to Ferromagnetism*, (McGraw-Hill, 1937) p. 153.

# The Reaction Rate Theory of Viscosity and Some of its Applications

By RAYMOND H. EWELL

*Department of Chemistry, Purdue University*

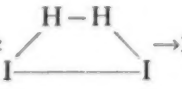
*Lafayette, Indiana*

THIS article will be concerned with the fundamental theory of the flow of matter which results in permanent deformation, and its application to several types of viscous and plastic materials. The general features of the theory apply to flow in all condensed phases, or more specifically to phases where viscous drag due to transfer of molecules between layers is negligibly small (i.e., gas type of viscous drag). This will include liquids of all sorts, glasses and resins, and crystals. About two years ago Eyring[1, 2] published an ingenious treatment of viscosity, plasticity and diffusion wherein flow was considered as a unimolecular chemical reaction in which the elementary process is the passing of a molecule (or other unit of flow) from one equilibrium position to another over a potential energy barrier. Since this treatment is based on the modern theory of chemical reaction rates as developed by Tolman, Kassel, Rice, London, Polanyi, Wigner, LaMer, Rodebush, Eyring and others (and reviewed by Eyring[3]), a brief discussion of this theory will be given first.

### *The Theory of Chemical Reaction Rates*

The most general treatment of reaction rates involves the concept of crossing a potential energy barrier by way of an activated complex

$$\text{reactants} \rightleftarrows \text{activated complex} \rightarrow \text{products}$$

e.g., $H_2 + I_2 \rightleftarrows$  $\rightarrow 2\,HI.$

The rate constant, $k_1$, for *a reaction of any order in any phase* is given by the equilibrium constant between the normal and activated states multiplied by the velocity at which the complexes cross the potential barrier and by the probability of their not returning. If the degree of freedom of the activated complex along which it approaches the potential barrier, crosses it and then flies to pieces, is taken as a translational one, we have:

$$\begin{aligned} k_1 &= \kappa K \frac{\bar{p}}{m} \\ &= \kappa K^* \frac{(2\pi mkT)^{\frac{1}{2}}}{h} \frac{\bar{p}}{m} \\ &= \kappa \frac{F_a^*}{F_n} e^{-\Delta E_a/kT} \frac{(2\pi mkT)^{\frac{1}{2}}}{h} \left(\frac{kT}{2\pi m}\right)^{\frac{1}{2}} \\ &= \kappa \frac{kT}{h} \frac{F_a^*}{F_n} e^{-\Delta E_a/kT}, \end{aligned} \tag{1}$$

where $K = \dfrac{\text{concentration of activated complex}}{\text{concentration of reactants}} = K^* \dfrac{(2\pi mkT)^{\frac{1}{2}}}{h}$,

*Potential energy barriers seem destined to play an ever increasing role in modern physics. In the present paper their use in the theory of rates of chemical reactions is described. The general theory of such reactions has been extended recently so that it is applicable to problems involving viscosity, plasticity and diffusion not only in liquids but also in glasses, resins and crystals.—Editor*

$$K^* = \frac{F_a^*}{F_n} e^{-\Delta E_a/kT}$$
$$= e^{\Delta S_a/k} e^{-\Delta E_a/kT}$$
$$= e^{-\Delta F_a/kT};$$

$m$ = mass of activated complex,
$\bar{p}$ = mean momentum of activated complex normal to the barrier,
$\kappa$ = transmission coefficient, the probability that once having crossed the barrier, the complex will not return, usually about unity,
$F_a^*$ = partition function of activated complex, omitting one translational degree of freedom,
$F_n$ = partition function of normal state,
$\Delta E_a$ = energy of activation,[a]
$\Delta F_a$ = free energy of activation,
$\Delta S_a$ = entropy of activation,
and other symbols have their usual meanings.

Equation (1) is the fundamental equation for any rate process, whether it be a chemical reaction, flow, or diffusion, i.e., for any process in which the slow step is passage over a potential barrier. This equation supersedes in generality the older idea of collision frequency, although Eyring has shown that they are equivalent for simple bimolecular reactions. Fig. 1 shows some of the energy relations in a rate process, and shows clearly that the rate of an exothermic reaction can have no relation to the heat or the free energy of the reaction. For practical purposes the contribution of the zero point energies to the activation energy can often be ignored.

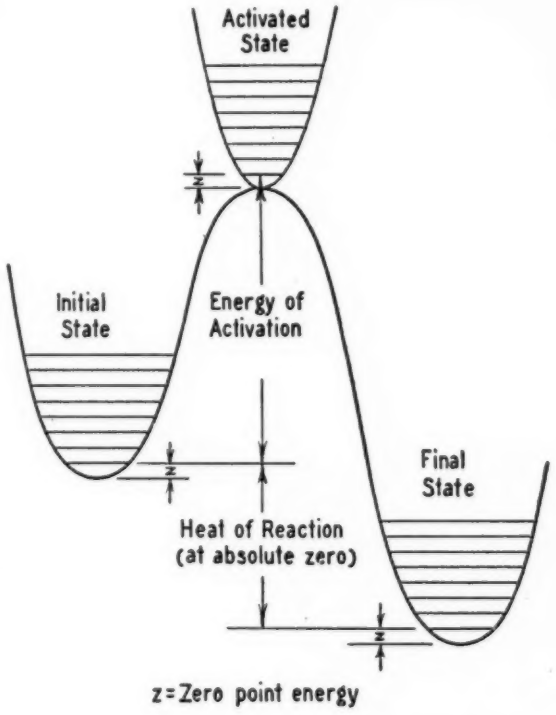



FIG. 1. Energy profile of the path of a reacting system.

[a] Throughout this article the same symbol will be used for energy per mole or per molecule, the particular one in question being distinguished by the value of the gas constant used with it.

### *Application of the Theory of Reaction Rates to Flow*

The development to be presented in this section follows quite closely that given by Eyring.[1] As stated in the introductory paragraph, flow will be considered as a unimolecular process in which the elementary process is the passing of a molecule from one equilibrium position to another over a potential barrier. The activation energy of this process takes the form of providing a hole for the molecule to flow into (or possibly of providing some extra space so that it may slip past its neighbors). If no velocity gradient exists in the liquid the energy profile of the flow process will be as shown by the full line in Fig. 2. A molecule will diffuse equally often from right to left as from left to right, and no net flow will occur. We will have for diffusion then

$$k_1 = \frac{kT}{h} \frac{F_a^*}{F_n} e^{-\Delta E_a/kT}.$$

This is Eq. (1) with $\kappa$, the transmission coefficient, omitted. Since $\kappa$ is usually about unity it will be omitted hereafter.

If, however, a velocity gradient is set up in the liquid by a force acting in the direction left to right in Fig. 2, there will be a decrease in the activation energy for a flow process in the direction left to right, and an equal increase in the direction right to left, so that the forward rate will be greater than the backward rate and a net flow results. Thus we have

$$k_f = \frac{kT}{h} \frac{F_a^*}{F_n} e^{-(\Delta E_a - \frac{1}{2} f\lambda_2\lambda_3\lambda)/kT}$$
$$= k_1 e^{f\lambda_2\lambda_3\lambda/2kT},$$
$$k_b = \frac{kT}{k} \frac{F_a^*}{F_n} e^{-(\Delta E_a + \frac{1}{2} f\lambda_2\lambda_3\lambda)/kT}$$
$$= k_1 e^{-f\lambda_2\lambda_3\lambda/2kT},$$

where $f$ = force per sq. cm tending to displace one layer with respect to another
$\lambda$ = distance between equilibrium positions in the direction of flow
$\lambda_2$ = distance between adjacent molecules in the direction of flow
$\lambda_3$ = distance between molecules in the plane of flow and normal to the direction of flow.

The area of a molecule in the plane of flow will be $\lambda_2\lambda_3$, the force acting on the molecule will then be $f\lambda_2\lambda_3$, and if the force acts through a distance $\lambda/2$, the energy of activation will be raised or lowered by $\frac{1}{2}f\lambda_2\lambda_3\lambda$. Fig. 2 shows these energy relations diagrammatically.

The viscosity is given hydrodynamically by

$$\eta=\frac{f\lambda_1}{\Delta V}=\frac{f\lambda_1}{\lambda(k_f-k_b)},$$

where $\lambda_1$ = perpendicular distance between adjacent layers of molecules,
$\Delta V$ = difference in velocity of two layers a distance $\lambda_1$ apart
$=\lambda(k_f-k_b)$,

so that

$$\eta=\frac{f\lambda_1}{\lambda k_1(e^{f\lambda_2\lambda_3\lambda/2kT}-e^{-f\lambda_2\lambda_3\lambda/2kT})}$$

$$=\frac{f\lambda_1}{\lambda k_1^2\sinh f\lambda_2\lambda_3\lambda/2kT}. \qquad (2)$$

The expansion of $\sinh x$ is $\sinh x=x+1/6x^3+1/120x^5+\cdots$, so that for ordinary viscous flow, where $f\lambda_2\lambda_3\lambda/2\ll kT$,

$$\sinh f\lambda_2\lambda_3\lambda/2kT\cong f\lambda_2\lambda_3\lambda/2kT$$

and we have

$$\eta=\frac{\lambda_1 h}{\lambda^2\lambda_2\lambda_3}\frac{F_n}{F_a^*}e^{\Delta E_a/kT}. \qquad (3)$$

Eq. (2) may be considered the fundamental equation of flow, and also Eq. (3) with the limitation mentioned.

Now if we assume that a molecule flows one intermolecular distance in each elementary process, i.e., that $\lambda=\lambda_2$, and for an equant molecule $\lambda_1=\lambda_2=\lambda_3$

$$\lambda_1/\lambda^2\lambda_2\lambda_3=N/V,$$

where $N$ = Avogadro's number
$V$ = molar volume.

If we assume that the degree of freedom corresponding to flow is a translational one, and that other degrees of freedom are the same for the initial and activated states:

$$\frac{F_n}{F_a^*}=\frac{((2\pi mkT)^{\frac{1}{2}}/h^3)V_f F_{\text{rot}}F_{\text{vib}}}{(2\pi mkT/h^2)V_f^{\frac{1}{3}}F_{\text{rot}}F_{\text{vib}}}=\frac{(2\pi mkT)^{\frac{1}{2}}}{h}V_f^{\frac{1}{3}},$$

where $V_f$ = the free volume per molecule. Eyring and Hirschfelder[4] have formulated the free volume

$$V_f^{\frac{1}{3}}=\frac{bRT}{V^{\frac{1}{3}}N^{\frac{1}{3}}(p+a/V^2)}\text{ per molecule}$$

$$=\frac{bRTV^{\frac{1}{3}}}{N^{\frac{1}{3}}\Delta E_{\text{vap}}}\quad\text{if}\quad\frac{a}{V^2}=\frac{\Delta E_{\text{vap}}}{V}\gg p,$$

where $\Delta E_{\text{vap}}$ = energy of vaporization per mole
$=\Delta H_{\text{vap}}-\Delta(pv)$,
$b=2$ for simple cubic packing and has not very different values for other types of packing.

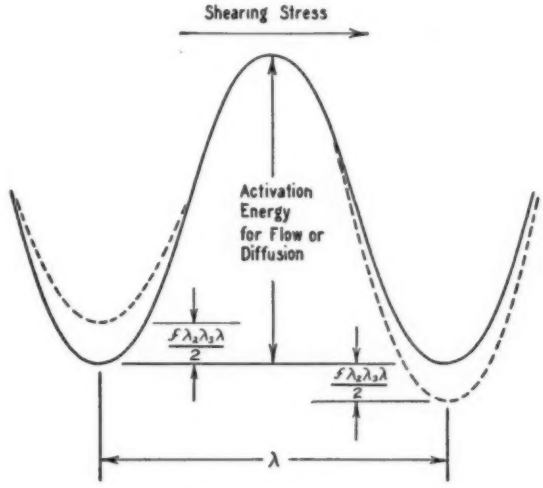



FIG. 2. Energy profile of the path of a flowing molecule.

As stated before the activation energy takes the form of providing a hole for the molecule to flow into, but this may not necessarily have to be a hole the full size of a molecule. A consideration of the attractive and repulsive forces involved in vaporization shows that for any liquid the energy

required to transfer a molecule to the vapor and leave a molecule-sized hole behind is $2\Delta E_{vap}$ (neglecting the $\Delta(pv)$ term), while the energy of either making a molecule-sized hole in a liquid, *or* of transferring a molecule from liquid to vapor without leaving a hole, is $\Delta E_{vap}$ (for further discussion of the theory of holes see reference 1). Thus the activation energy of the flow process will be some fraction of the energy of vaporization, $\Delta E_a = \Delta E_{vap}/n$.

Combining all these evaluations of the quantities of Eq. (3) we have

$$\eta = \frac{Nh}{V}\frac{(2\pi mkT)^{\frac{1}{2}}}{h}\frac{bRTV^{\frac{1}{3}}}{\Delta E_{vap}N^{\frac{1}{3}}}e^{\Delta E_{vap}/nRT}$$

$$= 1.090\cdot 10^{-3}\frac{M^{\frac{1}{2}}T^{\frac{1}{2}}}{V^{\frac{1}{3}}\Delta E_{vap}}e^{\Delta E_{vap}/nRT} \qquad (4)$$

if we take $b=2$, and $\Delta E_{vap}$ and $R$ are in calories per mole.

On testing Eq. (4) with carefully selected values of viscosity, heat of vaporization and density for 24 normal liquids, it was found that the experimentally observed temperature coefficient of viscosity was obtained when $n=3$ for liquids composed of spherically symmetrical molecules and $n=4$ for liquids composed of polar or elongated molecules. The results to the closest half-integer are given in Table I. As an example of the type of computation made for each liquid, Table II shows the calculations for carbon tetrachloride and Fig. 3 shows the plot of observed and calculated values.

The fact that there is a sharp separation into two classes of liquids (except $O_2$, $C_2H_4Cl_2$ and $C_2H_4Br_2$) with $n=3$ and $n=4$, respectively, is significant. The $n=3$ liquids all have approximately spherical fields of force about the molecules. This is certainly true of $CCl_4$, $CH_4$ and A, and is very nearly true of $N_2$ and CO when the probable shape of the potential energy shell at kinetic energy levels is considered. There is also considerable evidence that benzene and cyclohexane have approximately spherical fields of force. The fact that the activation energy for flow is just one-third the energy of vaporization for these spherical molecules leads to the interesting and useful concept that viscous flow is a sort of vaporization in one degree of freedom.

TABLE I.

| | |
|---|---|
| $n=3$ | $CCl_4$, $C_6H_6$, cyclohexane, $CH_4$, $N_2$, CO, A |
| $n=3\frac{1}{2}$ | $C_2H_4Cl_2$, $C_2H_4Br_2$, $O_2$* |
| $n=4$ | pentane, hexane, heptane, $CS_2$, $CHCl_3$, $C_6H_5CH_3$ ether, ethyl acetate, acetone, $C_2H_5I$, $C_2H_5Br$, $CH_3I$, $C_2H_4$ |

* The anomaly of $O_2$ as compared to $N_2$ and CO may be attributed to its paramagnetic property. No explanation is offered for the result for the ethylene halides.

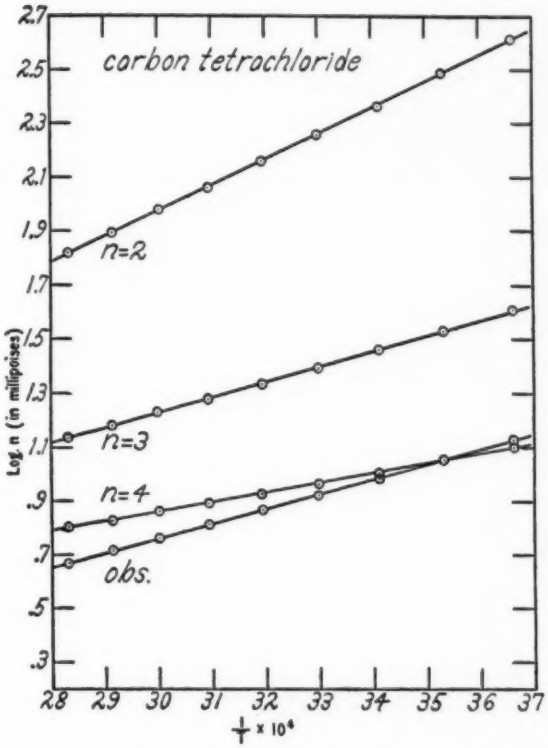



FIG. 3. Comparison of observed viscosities of carbon tetrachloride with those calculated from Eq. (4), using integral values of $n$.

The other group of liquids with $n=4$, however, are not of spherical symmetry and most of them are polar. For spherical molecules, where all orientations are equivalent, the activation energy of flow is the energy required to make a hole one-third the size of a molecule, i.e. $\Delta E_{vap}/3$, while for polar or elongated molecules where certain preferred orientations are possible a smaller fraction of the energy of vaporization is sufficient for activation. Naturally if there is a possibility of several orientations the molecule will assume that orientation during flow which requires the least activation energy. That this activation energy should be very close to one-fourth the

TABLE II. *Computation of viscosity of* $CCl_4$.

| | | | | | $n=2$ | | $n=3$ | | $n=4$ | |
|---|---|---|---|---|---|---|---|---|---|---|
| $t$°C | $T$°K | $V$ cc/ mole | $\Delta E_{vap}$ kcal./ mole | $\eta$(obs.) milli-poises | $\eta$(calc.) milli-poises | $\eta$(calc.)/$\eta$(obs.) | $\eta$(calc.) milli-poises | $\eta$(calc.)/$\eta$(obs.) | $\eta$(calc.) milli-poises | $\eta$(calc.)/$\eta$(obs.) |
| 0 | 273.1 | 94.3 | 7.56 | 13.47 | 413 | 30.7 | 40.3 | 2.99 | 12.7 | 0.95 |
| 10 | 283.1 | 95.4 | 7.43 | 11.33 | 305 | 26.9 | 33.8 | 2.98 | 11.3 | 1.01 |
| 20 | 293.1 | 96.6 | 7.30 | 9.69 | 232 | 24.0 | 28.8 | 2.98 | 10.1 | 1.04 |
| 30 | 303.1 | 97.8 | 7.18 | 8.42 | 182 | 21.6 | 24.8 | 2.95 | 9.2 | 1.08 |
| 40 | 313.1 | 99.0 | 7.06 | 7.38 | 145 | 19.7 | 21.7 | 2.94 | 8.5 | 1.13 |
| 50 | 323.1 | 100.3 | 6.93 | 6.53 | 115 | 17.6 | 19.1 | 2.92 | 7.8 | 1.19 |
| 60 | 333.1 | 101.6 | 6.81 | 5.84 | 96 | 16.4 | 17.1 | 2.93 | 7.3 | 1.24 |
| 70 | 343.1 | 103.0 | 6.68 | 5.24 | 78 | 14.9 | 15.3 | 2.92 | 6.7 | 1.28 |
| 80 | 353.1 | 104.4 | 6.56 | 4.68 | 66 | 14.2 | 13.8 | 2.95 | 6.4 | 1.36 |

energy of vaporization for such a wide variety of molecules is an interesting fact.

It will be noted that while the plot with $n=4$ in Fig. 3 is parallel to the observed, indicating the correct temperature coefficient, the two plots do not coincide. Viscosities calculated by Eq. (4) are always higher than the observed and the antilogarithm of the difference in ordinate of the parallel plots is the factor by which the calculated values are too large. For all $n=4$ liquids the calculated values are too large by a factor very nearly 2, while the $n=3$ liquids are too large by a somewhat larger and more variable factor varying from 2.5 to 3.5.

One possible explanation of this discrepancy may be that a flowing molecule may possess a "persistence of velocity" and when it has once acquired the necessary activation energy it may move more than one intermolecular distance, and $\lambda$ may be equal to $\lambda_2$, $2\lambda_2$, $3\lambda_2$, etc. for any individual elementary process. It would only be necessary for $\lambda$ to equal $2\lambda_2$ in 40 percent of the elementary processes to give a calculated viscosity too large by a factor of 2, or in 70 percent of the processes for a factor of 3.

Another possible explanation is that the flow process is very likely a bimolecular process rather than a unimolecular one. According to this mechanism two molecules in adjacent layers, which are moving relative to one another, rotate through an angle of about 90° as shown in Fig. 4.[5] The relative motion of successive layers in the liquid is the resultant of many such rotations of pairs of molecules. During the rotation the two molecules will sweep out a volume $\Delta V$, which is the extra volume necessary for the elementary process to take place, and is obviously $V/3$ and $V/4$ for the two classes of molecules.

The fact that the factors by which the calculated values are too large are about the same for all liquids, and the success of the theory in interpreting the temperature coefficient of vis-

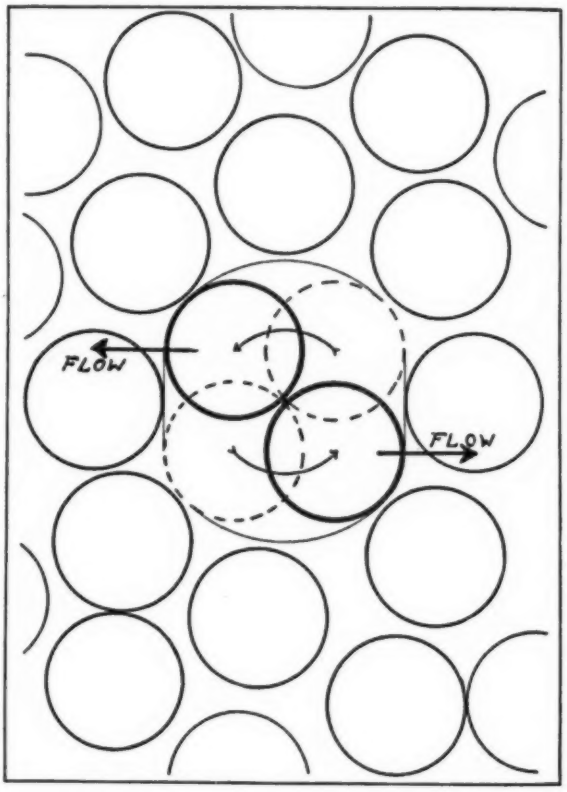



FIG. 4. Bimolecular mechanism of flow.*

* After Hirschfelder, Stevenson and Eyring, J. Chem. Phys. **5**, 907 (1937).

cosity show that the reaction rate theory is essentially correct, and that if the elementary process is a bimolecular one, it is essentially equivalent to the unimolecular mechanism, differing only by a small constant factor. Until this point in the theory is cleared up, viscosities calculated from Eq. (4) should be divided by 2

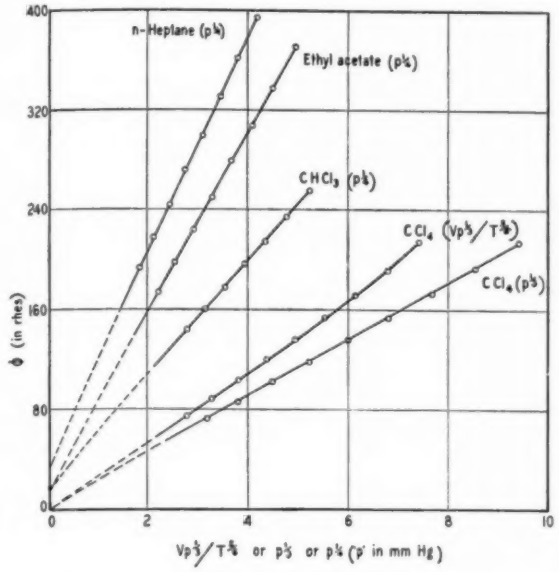



FIG. 5. The relation between fluidity and vapor pressure.

for $n=4$ liquids, and with somewhat more uncertainty by 3 for $n=3$ liquids.

### *The Relation Between Viscosity and Vapor Pressure*

Omitting the evaluation of the free volume, Eq. (4) would be

$$\eta=\frac{Nh}{V}\frac{(2\pi mkT)^{\frac{1}{2}}}{h}V_f^{\frac{1}{3}}e^{\Delta E_{vap}/3RT}$$

for a liquid composed of spherical molecules. Eyring and Hirschfelder[4] have evaluated the free volume of a liquid in terms of the vapor pressure as follows

$$V_f=(RT/p)e^{\Delta H_{vap}/RT}.$$

Substituting this in the above equation gives

$$\eta=\frac{N}{V}(2\pi mkT)^{\frac{1}{2}}\left(\frac{RT}{p}\right)^{\frac{1}{3}}e^{-(\Delta H_{vap}/3RT)}e^{\Delta E_{vap}/3RT}$$

$$=N\left(\frac{2\pi k}{N}\right)^{\frac{1}{2}}Re^{-\frac{1}{3}}\frac{M^{\frac{1}{2}}T^{5/6}}{Vp^{\frac{1}{3}}},$$

$$\text{or } \phi=\text{const. } Vp^{\frac{1}{3}}/T^{5/6}, \tag{5}$$

where $\phi$ = the fluidity, i.e., the reciprocal of viscosity. Fig. 5 shows $\phi$ plotted against $Vp^{\frac{1}{3}}/T^{5/6}$ for $CCl_4$, and it is disappointing that the curve is not a straight line. The plot is slightly curved, but the extrapolation appears to pass through the origin.

Figure 5 also shows $\phi$ plotted against $p^{\frac{1}{2}}$ for $CCl_4$, which gives a straight line passing through the origin. This is an interesting and significant relationship, even though empirical. It was also found that for nonspherical molecules, plots of $\phi$ *vs.* $p^{\frac{1}{2}}$ give straight lines, although the extrapolations do not pass exactly through the origin (see Fig. 5). However, plots of $\phi$ *vs.* $p^{\frac{1}{2}}$ for these latter liquids were strongly curved, emphasizing again the division into spherical and nonspherical molecules. This parallelism between fluidity and vapor pressure is fundamental to an understanding of the nature of flow.

### *The Variation of Viscosity with External Pressure*

In order to make a hole in a liquid under an external pressure, work must be done against both the internal and external pressures. The work against the internal pressure is $\Delta E_{vap}$, so that the general expression for the energy of making a molecule-sized hole in a liquid is $(\Delta E_{vap}+pV)$. At low pressures $pV$ can be neglected compared to $\Delta E_{vap}$, but in order to use Eq. (4) at more than a few atmospheres, $\Delta E_{vap}$ must be replaced by $(\Delta E_{vap}+pV)$ giving the more general equation

$$\eta=1.090\cdot10^{-3}\frac{M^{\frac{1}{2}}T^{\frac{1}{2}}}{V^{\frac{1}{3}}(\Delta E_{vap}+pV)}e^{(\Delta E_{vap}+pV)/nRT}. \tag{6}$$

The $\Delta E_{vap}$ in this equation is not the energy of vaporization measured at atmospheric pressure, for the energy of vaporization decreases as the volume decreases, finally becoming zero and then changing sign, so that at sufficiently high pressures the energy of transferring a molecule from the compressed liquid to the dilute vapor may actually become negative.

The energy of vaporization of a liquid under pressure has its most definite meaning in terms of the equation

$$\Delta E_{\text{vap}} = Vp_{\text{internal}} = V(\partial E/\partial V)_T$$
$$= V(T(\partial p/\partial T)_V - p) \quad (7)$$

due to Hildebrand.[6] If this is substituted in Eq. (6) we get

$$\eta = 1.090 \cdot 10^{-3} \frac{M^{\frac{1}{2}}T^{\frac{1}{2}}}{V^{5/3}(\partial p/\partial T)_V} \cdot \exp\left[V(\partial p/\partial T)_V/nR\right]. \quad (8)$$

Eqs. (6) and (8) are general equations for calculating the viscosity of a pure liquid at any temperature and pressure. Table III and Fig. 6

TABLE III. *Computation of the viscosity of ether as a function of pressure at 52.5°, using Eq. (8).*

| $p$ KG/CM² | $V$ CC/MOLE | $(\partial p/\partial T)_V$ KG/CM² / DEG. | $\eta$(CALC.) MILLI-POISES | $\eta$(OBS.) MILLI-POISES |
|---|---|---|---|---|
| 1 | 109.9 | 6.73 | 1.80 | 1.83 |
| 1000 | — | — | — | 3.61 |
| 2000 | 90.25 | 13.50 | 5.35 | 5.64 |
| 3000 | 86.70 | 16.00 | 8.00 | — |
| 4000 | 84.00 | 18.00 | 10.7 | 10.50 |
| 5000 | 81.80 | 19.80 | 14.0 | — |
| 6000 | 80.05 | 21.40 | 17.7 | 17.58 |
| 7000 | 78.50 | 22.85 | 22.0 | — |
| 8000 | 77.15 | 24.20 | 26.3 | 27.75 |
| 9000 | 76.05 | 25.40 | 30.9 | — |
| 10000 | 75.05 | 26.50 | 35.8 | 42.69 |
| 11000 | 74.10 | 27.50 | 40.8 | — |
| 12000 | 73.25 | — | — | 64.24 |

show a test of Eq. (8) for ether at 52.5°, using $n=4$ and dividing the calculated values by 2 as mentioned above in the test of Eq. (4). $V$ and $(\partial p/\partial T)_V$ were evaluated from the PVT data of Bridgman.[7] The observed viscosities under pressure were also taken from Bridgman.[8] It is seen that the equation fails above about 7000 kg/cm². This failure may be due to the failure of Eq. (7) or to any of several other causes which need not be discussed here, but the close agreement up to 7000 kg/cm² gives striking confirmation to the fundamental correctness of the theory.

Equation (6) will give *approximate* results for pressures below 2000 kg/cm² when the ordinary energy of vaporization measured at atmospheric pressure is used, thus requiring PV data only at the temperature in question. This is possible because the internal pressure and energy of vaporization change very little for pressures up to about 2000 kg/cm². Eq. (6) can also be used to calculate the internal pressures of liquids if experimental values of the viscosity under pressure are available—just the reverse of the calculation in the last paragraph. The original paper[2] illustrates both these applications.

An obvious application of Eq. (8) is to the calculation of the viscosity of lubricants under high pressures. The fact that the viscosity of liquids increases rapidly with pressure is, of course, of great importance in many lubrication problems. Mineral lubricating oils are mixtures of many molecular species, and since Eq. (8) is true only for pure liquids, no attempt will be made here to calculate the absolute viscosities of oils under pressure. However, a calculation of the relative viscosities at two different pressures seems more likely to succeed. Eq. (8) gives

$$\frac{\eta_{p_2}}{\eta_{p_1}} = \frac{(V_r^{5/3}(\partial p/\partial T)_V)_{p_1}}{(V_r^{5/3}(\partial p/\partial T)_V)_{p_2}} \cdot \exp\frac{M}{nRd_0}\left[\left(V_r\left(\frac{\partial p}{\partial T}\right)_V\right)_{p_2} - \left(V_r\left(\frac{\partial p}{\partial T}\right)_V\right)_{p_1}\right],$$

where $V_r$ = relative volume,
$d_0$ = reference density,
$n=4$,

and $R$ must be in the appropriate energy unit.

Table IV shows the results of testing this equation. $(\partial p/\partial T)_V$ and $V_r$ were evaluated from the PVT data of Dow[9] on a Pennsylvania oil, and the observed viscosities were interpolated from the data of Dow[10] on *another* Pennsylvania oil, whose molecular weight was 485±10 percent. The value of the molecular weight in such a calculation is the most uncertain factor. The

TABLE IV. *Computation of the viscosity of Pennsylvania oil as a function of pressure at 57.5°C. $M=485\pm10\%$. $d_0=0.8524$.*

| $p$ KG/CM² | $V_r$ | $(\partial p/\partial T)_V$ KG/CM² / DEG. | $\frac{\eta_{1500}}{\eta_{375}}$(CALC.) | $\eta$(OBS.) POISES | $\frac{\eta_{1500}}{\eta_{375}}$(OBS.) |
|---|---|---|---|---|---|
| 375 | 0.9875 | 10.81 | 6.96 | 0.77 | 6.95 |
| 1500 | 0.9412 | 12.78 | | 5.35 | |

TABLE V.

| LIQUID | $\Delta E_{vis}$ CAL./MOLE | $\Delta E_{vap}$ AT B.P. CAL./MOLE | $\frac{\Delta E_{vap}}{\Delta E_{vis}}$ | $n$ |
|---|---|---|---|---|
| $CCl_4$ | 2500 | 6600 | 2.66 | 3 |
| $C_6H_6$ | 2540 | 6660 | 2.62 | 3 |
| Cyclohexane | 2890 | 6700 | 2.32 | 3 |
| $CH_4$ | 719 | 1820 | 2.53 | 3 |
| A | 516 | 1420 | 2.75 | 3 |
| $N_2$ | 449 | 1210 | 2.70 | 3 |
| CO | 466 | 1310 | 2.80 | 3 |
| $O_2$ | 398 | 1470 | 3.69 | $3\frac{1}{2}$ |
| $C_2H_4Cl_2$ | 2270 | 6930 | 3.05 | $3\frac{1}{2}$ |
| $C_2H_4Br_2$ | 2590 | 7890 | 3.04 | $3\frac{1}{2}$ |
| Pentane | 1580 | 5510 | 3.50 | 4 |
| Hexane | 1715 | 6220 | 3.61 | 4 |
| $CHCl_3$ | 1760 | 6630 | 3.76 | 4 |
| $C_2H_5I$ | 1720 | 6400 | 3.72 | 4 |
| $C_2H_5Br$ | 1585 | 6080 | 3.84 | 4 |
| $CS_2$ | 1280 | 5920 | 4.63 | 4 |
| $C_6H_5CH_3$ | 2120 | 7240 | 3.42 | 4 |
| Ether | 1610 | 5700 | 3.54 | 4 |
| Acetone | 1655 | 6400 | 3.86 | 4 |
| $C_2H_4$ | 793 | 3500 | 4.41 | 4 |

calculation was made using $M=440$, and the good agreement is due to this value having been found by trial beforehand. It is, however, within the limit of uncertainty of the experimentally determined molecular weight (of the oil on which the viscosity measurements were made). Considering the fact that the data were from two different Pennsylvania oils, the results appear promising, at least. The pressures of 375 and 1500 kg/cm² were not arbitrarily selected, but were the lowest and highest pressures, respectively, at which $(\partial p/\partial T)_V$ could be evaluated from the PVT data. Experimental data suitable for a test over a wider range of pressures would be desirable, since Eq. (8) appears to hold up to 7000 kg/cm² for pure liquids.

### *Comparison of the Theory with the Empirical Equation* $\eta = Ae^{B/T}$

It is a well-known experimental fact that a plot of $\log \eta$ *vs.* $1/T$ is a straight line for all normal liquids, or that

$$\eta = Ae^{B/T} = Ae^{\Delta E_{vis}/RT} \qquad (9)$$

and this relation is now one of the best established empirical rules regarding liquids. All normal liquids, including hydrocarbons and their halogen derivatives, ketones, ethers, sulfides, esters, acid anhydrides, liquefied gases, covalent inorganic compounds, etc. obey the equation within experimental error, and also most liquid metals and fused salts. In fact, it is so universally obeyed by normal liquids that it can be considered empirically as a necessary condition for a liquid of unchanging molecular state, although it cannot be considered a sufficient condition. Fig. 7 shows a plot of some liquids representing several types of compounds.

This relation has had a long and devious history, having been independently "rediscovered" periodically ever since 1913. Following is a chronological summary of the independent publications, the asterisk indicating those which can be thrown into the identical form of $\eta = Ae^{B/T}$:

| | | |
|---|---|---|
| 1913 | de Guzman[11] | * |
| 1916 | Arrhenius[12] | $\eta v^{\frac{1}{3}} = Ae^{B/T}$ |
| 1917 | Kendall and Monroe[13] | * |
| 1918 | Drucker[14] | * |
| 1923 | Raman[15] | * |
| 1925 | Fulcher[16] | $\eta = Ae^{B/T+C}$ |
| 1926 | Dunn[17] | * |
| 1929 | Busse and Karrer[17a] | * |
| 1930 | Andrade[18] | * |
| 1930 | Sheppard[19] | * |
| 1931 | Tonomura[20] | $\eta = Ae^{B/T+C}$ |
| 1933 | Cragoe[21] | $\eta v = Ae^{B/T}$ |
| 1934 | Andrade[22] | $\eta v^{\frac{1}{3}} = Ae^{B/vT}$ |
| 1937 | Souders[23] | $\eta v = Ae^{B/T}$ |

The symbol $V$ is the specific volume.

The variant forms which contain a volume factor make little or no improvement over the simple equation. In the recent paper by Souders he shows that there is little to choose between his equation and Andrade's 1934 equation, except that the former is easier to use. That $\eta = Ae^{B/T}$ and the variant forms containing the volume work about equally well is due to the slow variation of the volume with temperature.

The several deductions of this relationship have all been based on loose thermodynamic analogies with the vapor pressure, or on vague over-simplified applications of the Boltzmann factor, or on wholly empirical grounds, so that the equation to date must be considered as being purely an empirical one.

From the work of Eyring and the writer[1, 2] it is obvious that the constant $\Delta E_{vis}(=RB)$ is connected with the energy of vaporization. De

Guzman noticed that in a few cases $\Delta E_{vis}$ was nearly equal to the heat of fusion, but that this was not true in many more cases. Later writers

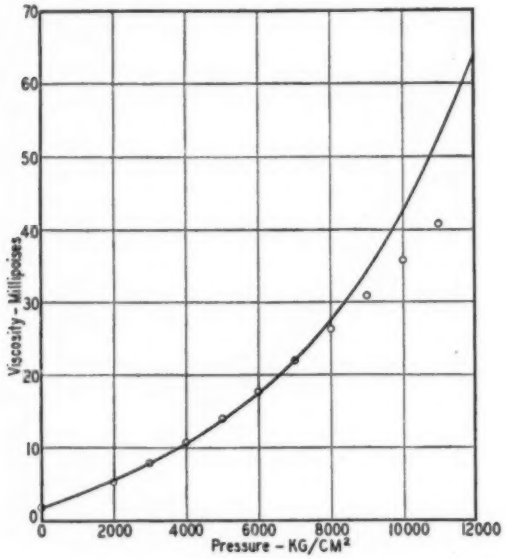



FIG. 6. Calculated and observed viscosities of ether as a function of the external pressure.

have persisted in continuing this heat of fusion relationship, although its empirical basis is a meager coincidence and its theoretical basis is nonexistent. Various terms have been applied to $\Delta E_{vis}$, such as "heat of disorientation," "heat of de-aggregation," "energy required to overcome the attraction of its neighbors," *et al.* These terms are all quite vague, but it is difficult to see how they could have led their authors so uniformly to the erroneous relationship with the heat of fusion. The adoption of the heat of fusion must have been intuitional in any case since until November 1937[5] there had been no adequate theory of the heat of fusion and the melting process.

Equation (4) can be put in the form

$$\eta = Nh/Ve^{-\Delta S_a/R}e^{\Delta E_a/RT},$$

$$\eta V = Nhe^{-\Delta F_a/RT}, \qquad (10)$$

where $\Delta F_a = \Delta E_a - T\Delta S_a$

$$= \Delta E_{vap/n} - T\Delta S_a,$$

$\Delta S_a$ is a small negative number for most liquids ranging from $-1$ to $-5$ cal./deg., so that $-T\Delta S_a$ is a positive quantity of the order of 1000 cal. $\Delta E_a$ decreases and $-T\Delta S_a$ increases with increasing temperature so that $\Delta F_a$ is more independent of temperature than either $\Delta E_a$ or $T\Delta S_a$.

The form of Eq. (10) is that of the variants of Cragoe and of Souders, where these latter have the specific volume and Eq. (10) has the molar volume. Evidently $\Delta E_{vis}$ is to be identified with the free energy of activation of flow, and the constant $A$ should be equal to $Nh$. $Nh$ has the value 1/250, while the experimentally determined constant $A$ is about 1/100 for most liquids. This simply means that the free energy is not entirely independent of temperature, but is of the form $\Delta E_{vis} + bT$, where $b$ would have a value of about 2 to give agreement between the theory and experiment. The experimental constant $\Delta E_{vis}$ is in fact then the constant part of the free energy of activation for flow.

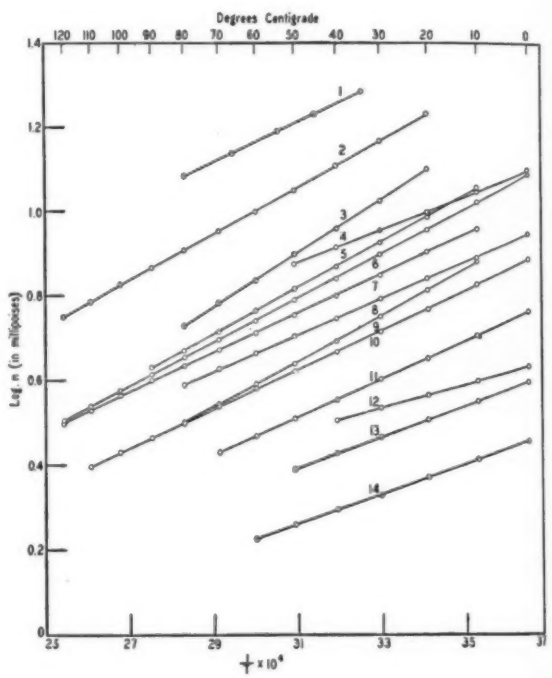



FIG. 7. Plots of log $\eta$ *vs.* $1/T$ for some normal liquids.†

† The data were taken from Landolt-Börnstein *Tabellen*, and are largely those of Thorpe and Rodger. The liquids represented are:

1. Phosphorus
2. Ethylene dibromide
3. Dioxane
4. Bromine
5. Carbon tetrachloride
6. Acetic anhydride
7. Chlorobenzene
8. Isopropyl iodide
9. Benzene
10. Toluene
11. Ethyl acetate
12. Carbon disulfide
13. Acetone
14. Diethyl ether

### *The Empirical Criterion* $\Delta E_{vap}/\Delta E_{vis}$

If Eq. (4) is differentiated with respect to $1/T$ we get

$$\Delta E_{vis}{}^{b} = Rd \ln \eta/d(1/T)$$

$$= \frac{\Delta E_{vap}}{n} - \tfrac{3}{2}RT + \tfrac{2}{3}\alpha RT^2 + \frac{\Delta C_V}{\Delta E_{vap}} RT^2,$$

where $\alpha$ = coefficient of thermal expansion.

$\Delta C_V$ = change in molal heat capacity on vaporization.

The first two terms decrease and the last two increase with rising temperature such that $\Delta E_{vis}$ remains constant. The last two terms together are about equal in magnitude to the second term, and so $\Delta E_{vis}$ should be approximately equal to $\Delta E_{vap/n}$, or $n$ should be equal to $\Delta E_{vap}/\Delta E_{vis}$. Table V gives a comparison of $n$ taken from Table I with the ratio $\Delta E_{vap}/\Delta E_{vis}$ at the boiling point for all the liquids in the table.

It is seen that for all the $n=3$ liquids $\Delta E_{vap}/\Delta E_{vis}$ ranges from 2.3 to 2.8, and for all the $n=4$ liquids from 3.4 to 3.9 (except $CS_2$ and $C_2H_4$ which are still higher). If the comparison had been made at a lower corresponding temperature, e.g., about four-fifths the absolute boiling point, all the values of the ratio would have been higher (since $\Delta E_{vap}$ would be larger), and the values of the ratio would have clustered around 3 for the $n=3$ liquids and around 4 for the $n=4$ liquids (except $C_2H_4$ and $CS_2$). These results show that even though $\Delta E_{vap}$ varies slowly with the temperature, the value of the ratio $\Delta E_{vap}/\Delta E_{vis}$ at any temperature may be taken as a rough index of the size and shape of the molecule, or more precisely, of the unit of flow. If $\Delta E_{vap}/\Delta E_{vis}$ is much larger than 3 or 4, it indicates that the unit of flow is much smaller than the unit of vaporization (which can be taken as the single molecule in most cases), and if $\Delta E_{vap}/\Delta E_{vis}$ is much smaller than 3 or 4, the reverse is indicated. Examples of both these cases will be discussed below.

[b] $\Delta E_{vis}$ is the experimental quantity $Rd \ln \eta/d(1/T)$ and will be called the viscosity energy. This is a constant independent of temperature for most liquids, but decreases with rising temperature for liquids whose structure changes with temperature.

While theoretically $\Delta E_{vis}$ is to be identified with the constant part of the *free energy* of activation for flow, this discussion has shown that it is approximately equal to the *energy* of activation for flow. While this relation must be considered as empirical until the theory is developed further, it will be found to be a very useful one.

The empirical criterion $\Delta E_{vap}/\Delta E_{vis}$ was applied by Ewell and Eyring[2] to several types of liquids with results as follows:

(a) *Liquid metals.* Most metals give a linear plot of log $\eta$ *vs.* $1/T$ just as normal covalent liquids do. The most striking fact regarding the metals is the large value of the $\Delta E_{vap}/\Delta E_{vis}$ ratio, which ranges from 8 to 25 for different metals as compared to 3 or 4 for normal liquids. This means that the activation energy for flow is much less than a third the energy of vaporization and that the unit of flow is much smaller than the unit of vaporization. The unit of vaporization is the atom, of course, and the unit

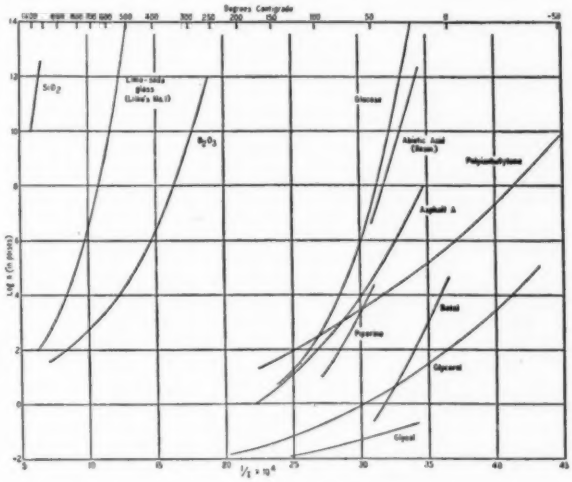



FIG. 8. Plots of log $\eta$ *vs.* $1/T$ for some associated, very viscous and glass-forming liquids.‡

‡ The data were from the following sources: $SiO_2$, Volarovich and Leontievna, J. Soc. Glass Tech. **20**, 139 (1936); soda lime glass, Lillie, J. Am. Cer. Soc. **16**, 619 (1933); $B_2O_3$, Parks and Spaght, Physics **6**, 69 (1935); glucose, Parks, Barton, Spaght and Richardson, ibid. **5**, 193 (1934); abietic acid, Bingham and Stephens, ibid. **5**, 217 (1934); polyisobutylene, Ferry and Parks, ibid. **6**, 356 (1935); asphalt, Pittman and Traxler, ibid. **5**, 221 (1934); glycol, Bingham, and Fornwalt J. Rheol. **1**, 373 (1930); piperine, betol and glycerol (below 0°C), Tammann and Hesse, Zeits. f. anorg. allgem. Chemie **156**, 245 (1926); glycerol (above 0°C), unpublished data by the writer and R. F. Newton.

of flow must then be the much smaller metal ion, i.e., the atom free of its valence electrons.

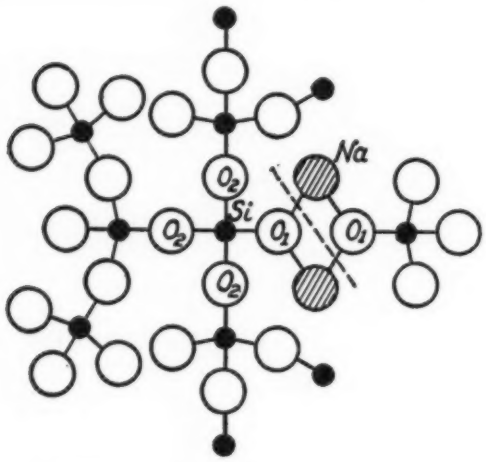



FIG. 9. Schematic representation of the structure of a soda-silica glass.§

On this hypothesis the activation energy of flow would be

$$\Delta E_{vis}=\frac{\Delta E_{vap}}{3}\times\frac{\text{volume of ion}}{\text{volume of atom}}$$

$$=\frac{\Delta E_{vap}}{3}\times\left(\frac{\text{radius of ion}}{\text{radius of atom}}\right)^3$$

or

$$\frac{\Delta E_{vap}}{\Delta E_{vis}}\times\left(\frac{\text{radius of ion}}{\text{radius of atom}}\right)^3=3.$$

A test of this relation for nine liquid metals showed it to be true within the experimental error of the quantities involved. While this result is interesting and suggestive, the hypothesis is undoubtedly too simplified and should not be taken too literally. The qualitative concept of the metal ions flowing free of their valence electrons is, of course, quite in harmony with modern ideas on the structure of metals.

(b) *Liquid sulfur.* As is well known, liquid sulfur is a fluid yellow liquid between the melting point and about 160° and also in the supercooled liquid region below the melting point. Above 160° the viscosity increases rapidly with rising temperature, increasing several thousand-fold between 160° and 190°, and thereafter the viscosity decreases in a normal way. It is thought that below 160° the liquid is composed of $S_8$ rings ($S_\lambda$), and at higher temperatures of $S_n$ chains ($S_\mu$), at least in part. The plot of log $\eta$ *vs.* $1/T$ gives two linear portions, one below 160° and one above 250°, whose slopes give $\Delta E_{vis}$ equal to 7.04 and 18.35 kcal., respectively. A comparison of these values with the known heat of vaporization led to a confirmation of the $S_8$ ring structure in the lower temperature range and to an $S_{36}$ chain formula at higher temperatures. The exact interpretation of the figure $S_{36}$ is not certain; the liquid may be an equilibrium mixture of $S_8$ rings and chain molecules of different lengths, or of the chain molecules only, so as to give an average of $S_{36}$ as deduced from the experimental value of $\Delta E_{vis}$.

(c) *Long chain hydrocarbon.* It has previously been shown that linear molecules require less than $\frac{1}{3}$ the energy of vaporization for activation of the flow process. If the molecules in a series such as the normal paraffins all flowed in the extended form, it would seem that they would require a smaller and smaller fraction of the energy of vaporization as the length of the chain increased. The data showed that $\Delta E_{vap}/\Delta E_{vis}$ was constant at 4 for all the normal paraffins from $C_5H_{12}$ to $C_{18}H_{38}$, and this led to the conclusion that all these hydrocarbons were curled into a ball shape during the flow process. Such a concept is in harmony with Langmuir's conclusion that molecules are curled in this way in the vapor state, since flow can be thought of as vaporization in one degree of freedom This suggests that the liquid hydrocarbon contains both extended and curled molecules, and the activated molecules which are responsible for flow, diffusion and vaporization of the liquid are of the curled type. For further details the reader is referred to the original paper.

### *General Remarks on the Application of the Theory to Other Than Normal Liquids*

From the point of view of viscous behavior and of most other physical properties, liquids may be classified as follows:

I. Intermolecular forces are exclusively undirected forces (van der Waals, dipole and repulsive forces)

§ After B. E. Warren and A. D. Loring, J. Am. Cer. Soc. **18**, 272 (1935).

(a) Relatively small molecules
E.g. $CCl_4$, $O_2$, $Cl_2$, $C_2H_5Br$, ether, decane, acetone, ethyl acetate, etc.
(b) Very long chain molecules
E.g., Linear polymeric resins, such as polystyrene, polyisobutylene, polyvinyl esters, etc.
Very long chain hydrocarbons, such as sone lubricating oils
Selenium and $\mu$-sulfur
Raw rubber

II. Cohesive forces are directed forces in part
(a) Hydrogen bond structure
E.g., $H_2O$, $NH_3$, alcohols, sugars, alkaloids, nonhardening resins such as cresol-formaldehyde: in general, any liquid whose molecules have OH or NH groups.
(b) Covalent bond structure
E.g., $SiO_2$, $GeO_2$, $B_2O_3$, $P_2O_5$, $BeF_2$ and all silicate, borate and phosphate liquids which are not too basic.
Hardening resins such as phenol-formaldehyde, glycerol-phthalic anhydride, etc.

III. Metallic liquids—Molten metals

IV. Ionic liquids—Molten salts.

The applications of the theory thus far have been limited to Group I(a), the so-called normal liquids (and also to liquid metals and $\mu$-sulfur). Glass-forming liquids are limited to Groups I(b), II(a) and II(b), and in fact most liquids in these groups do form glasses. Groups III and IV are really subdivisions of Group I, but are classified separately for obvious reasons.

In applying the theory to other than normal liquids three points are of importance:

(1) What is the structure of the liquid?
(2) What is the unit of flow?
(3) What is the energy of vaporization of this unit of flow?

In normal liquids the structure of the liquid is fairly certainly known and the unit of flow is without question the molecule, and furthermore the unit of flow and the unit of vaporization are undoubtedly the same. In other than normal liquids the unit of flow, the unit of vaporization, and the stoichiometric molecule need not necessarily be the same.

For any liquid we have the generalization that the energy of making a molecule-sized hole in the liquid is equal to the energy of vaporization, and that

$$\Delta E_{vap} = \tfrac{1}{2}\sum_{\text{forces}} \text{energy of all attractive and repulsive forces}$$

$$= \tfrac{1}{2}\Big[\sum_i E_{1i}\,(\text{van der Waals}) + \sum_i E_{1i}\,(\text{dipole}) + \sum_i E_i\,(\text{structural bonds})\Big], \quad (11)$$

neglecting repulsive forces. When an elementary flow process takes place in a liquid possessing a directed bond structure, not only must a hole be provided, but in addition structural bonds must be broken. It seems probable that the energy of activation for flow in a structural liquid will then be $\frac{1}{3}$ or $\frac{1}{4}$ of the energy of vaporization due to undirected forces *plus all* of the energy of vaporization due to structural bonds:

$$\Delta E_{vis} = 1/n\Delta E_{vap}(\text{undirected forces}) + \Delta E_{vap}(\text{structural bonds}). \quad (12)$$

In normal liquids the last term is zero. In liquids containing hydrogen bonds (Group II(a)) the two components of the energy of vaporization will be comparable in magnitude since the energy of a hydrogen bond is about 6000 cal. But in Group II(b) liquids the energy of breaking the structural bonds will be practically the whole energy of activation, and van der Waals and dipole attractions can be neglected in comparison, e.g., liquid $SiO_2$.

All this applies, of course, to the actual unit of flow, whatever it may be. The unit (or units) of flow can usually be determined if the structure of the liquid and the strength of all the bonds concerned are known. Flow will take place by all possible units of flow, but chiefly by the most economical one, i.e., the one with the smallest energy of activation. For example in liquid $SiO_2$ the units of flow SiO, $SiO_2$, $SiO_3$, and $SiO_4$ would all require the breaking of 4 Si—O bonds, while any unit of flow containing two silicon atoms would require the breaking of 6 Si—O bonds. Since the energy of an Si—O bond is 100,000 cal., the unit of flow will be $SiO_2$, or an appropriate mixture of all four of the groups

mentioned above. Whatever the units of flow may be, the summation of the flow processes must, of course, add up to the gross composition of the liquid.

The formulation of the entropy of activation of flow in these complex liquids is generally impossible so that the best that can be done is to assume that it is the same at different temperatures and pressures for the same liquid, or that it is the same at different compositions of a liquid mixture, such as a series of $Na_2O-SiO_2$ glasses. For instance, in the calculation of the change of viscosity of a lubricating oil with pressure it was assumed that the entropy of activation was the same at different pressures. These are only approximations, but often give useful results. In any case the entropy factor is not as important as the energy factor, and a high energy of activation will always mean a high viscosity, and *vice versa*.

The degree of structure in the liquids of Group II changes with the temperature. The higher the temperature the fewer bonds are present in the liquid and the weaker these bonds are. As a result of one or both these factors the energy of activation of flow for such liquids becomes less and less with increasing temperature. Fig. 8 shows plots of $\log \eta$ *vs.* $1/T$ for some liquids representing Groups I(b), II(a) and II(b). The curves are all practically straight at viscosities greater than $10^9$ poises, and this means that in this range the bond structure of these liquids has become constant and probably that the maximum possible number of bonds have been formed, e.g., four Si−O bonds per $SiO_2$ in silica glass.

### *Glasses and Resins*

Glasses and resins are grouped together because they are in reality identical in the basis of their properties and in their *raisons d'etre*. Glasses and resins are both found in Groups I(b) (e.g., selenium and polyisobutylene), II(a) (e.g., glucose and cresol-formaldehyde resin), and II(b) (e.g., silica and glyptal), and for the same reasons. The liquids of Group II readily supercool and vitrify since the energy of orienting a molecule preparatory to crystallization involves the breaking of structural bonds. Hence the energy of activation for the crystallization process is very high, and the rate of crystallization is extremely slow.[c] Since hydrogen bonds are relatively weak (about 6000 cal.), the tendency to vitrify is not as strong in Group II(a) as in II(b).

The Group I(b) liquids usually vitrify, and the reason for this is usually stated to be that the tangled "brush heap" structure of the long chains prevents crystallization. More precisely stated, the orientation required for crystallization involves a large decrease in entropy for these long chains so that the entropy of activation for the crystallization process is a large

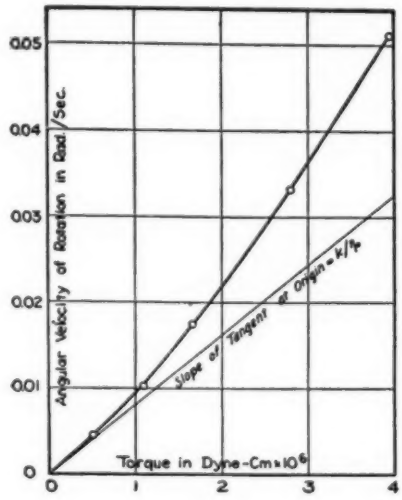



FIG. 10. "Consistency" curve for polyisobutylene at 275.8°K.‖

negative number and $-T\Delta S$ is a large positive number, and the rate of crystallization is very slow, even though the energy of activation may not be very large. In addition, these long chain glass-forming liquids are composed of chains of varying lengths, possibly several hundred different molecular species, as in the polystyrene resins, and it may be that there is never a stable crystalline phase, until temperatures corresponding to extremely high viscosities are reached. The same may be true for mixtures of hydrocarbons such as lubricating oils.

---

[c] The rate of crystallization has, of course, no relation to the magnitude of the thermodynamic free energy change on crystallization.

‖ After J. D. Ferry and G. S. Parks, Physics, **6**, 359 (1935).

### *The Viscosity of Silicate Glasses*

The structures of several inorganic glasses have been studied, principally by Warren and his co-workers. Silica glass[24] has a random structure with each silicon surrounded by four oxygens tetrahedrally at a distance of 1.60A. It is obvious that whatever the unit (or units) of flow may be, four Si−O bonds will have to be broken per unit of flow. The energy of an Si−O bond is known to be about 100,000 cal. from thermal data, and so $\Delta E_{vis}$ should be a half of four times this energy, or 200,000 cal. at any temperature at which this four-coordinated structure is complete. The slope of the curve in Fig. 8 multiplied by $2.303R$ gives $\Delta E_{vis} = 175{,}000$ cal.

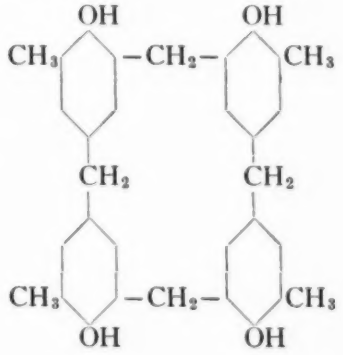
The addition of metal oxides to $SiO_2$ greatly reduces the viscosity. Soda-silica glasses[25] have a structure in which some of the Si−O−Si bridges are replaced by the electrostatically bound $ONa_2O$ bridges shown diagrammatically in Fig. 9, taken from Warren and Loring's paper.[25] The sodium and $O_1$ oxygens are univalent ions and the interionic distances are believed to be the following: $Na-O_1 = 2.35$, $Na-Na = 3.85$, and $O_1-O_1 = 2.65$A. Four attractions and two repulsions, all electrostatic, are involved in this bridge, and a simple calculation shows that the bridge breaks most easily into two ONa groups as indicated by the dotted line in Fig. 9 and that this requires 70,000 cal. Whatever the unit of flow may be, it will certainly involve four bonds. If three of these are Si−O bonds and one is an $ONa_2O$ bridge, the activation energy will be 185,000 cal.; if two are Si−O bonds and two are $ONa_2O$ bridges, the activation energy will be 170,000 cal., etc. The higher the soda content the lower will be the activation energy and the lower the viscosity. The change of viscosity with composition in $Na_2O-SiO_2$ glasses can be accounted for roughly quantitatively in this way.

Similar considerations apply to other binary glasses, such as $PbO-SiO_2$ glasses, which have a replacement of Si−O−Si bridges by Si−O−Pb−O−Si bridges.[26] The energy required to break this bridge is probably about 40,000 cal. The viscosity measurements of Bair[26] on $Pb-SiO_2$ glasses and of Taylor and Dear[27] on $Na_2O-SiO_2$ glasses show that PbO is much more potent in reducing the viscosity of silicate glasses than $Na_2O$ is.

### *The Viscosity of Resins*

Resins which have (2,3) or greater reactivity, such as phenol-formaldehyde, glycerol-phthalic anhydride, etc., are called hardening, convertible or infusible resins. When completely polymerized, such resins form a three-dimensional network of covalent bonds and are entirely analogous to silica glass in every respect. Flow in such a resin involves breaking several C−C or C−O bonds, each of which have energies of about 80,000 cal. This makes the activation energy for flow nearly as great as that of silica glass, and no softening would occur below a red heat at the lowest. Naturally the resin decomposes before that temperature is ever reached, and so they are called hardening or infusible resins.

Resins such as cresol-formaldehyde and some others which have only (2,2) reactivity are presumed to form complexes of limited size, but a three-dimensional network is impossible. The following diagram




illustrates the type of structure, although the complexes in a resin will be much larger. Such resins are called nonhardening, nonconvertible or fusible resins. Flow in such a resin involves breaking a number of hydrogen bonds, (at 6000 cal. each) formed between the complexes through the OH groups, plus a fairly large energy of activation due to van der Waals and dipole forces for such large molecules. Such a resin will therefore be very viscous, even hard and glassy at low enough temperatures, depending on the size of the complexes, but its energy of activation for flow will be of an entirely different order of magnitude from that of the (3,2) reactive resins,

and they would be expected to soften and become relatively fluid at temperatures lower than the decomposition temperature.

Linear polymeric resins, such as polystyrene, polyisobutylene, polyvinyl esters, etc., also have (2,2) reactivity and are usually included in the nonhardening, nonconvertible, fusible class. Flow in such a resin involves only energy of activation due to van der Waals and dipole forces since

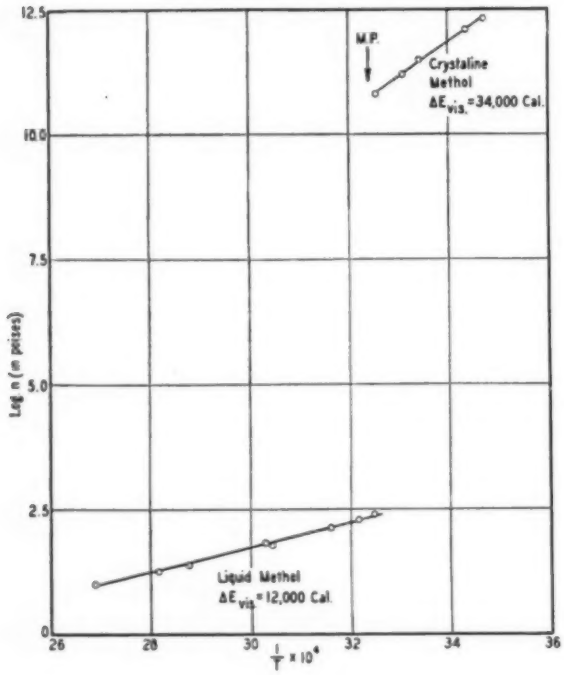



FIG. 11. Plot of log $\eta$ *vs.* $1/T$ for liquid and crystalline menthol.

*Note:* The crystal viscosities were plotted as poises and the liquid viscosities as millipoises so the two portions of the plot should be separated by 3 more units of ordinate than appears in the figure. Also, methol should read menthol.

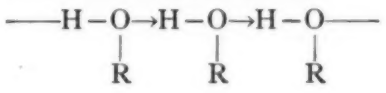
there are no structural bonds between the linear complexes. The energy of vaporization of a linearly polymerized molecule would be approximately $n$ times the energy of vaporization of one of the unit molecules in the polymerized molecule. The unit of flow in such a polymer will undoubtedly be a whole polymerized molecule, since breaking one of the covalent bonds within the molecule doesn't seem likely, but it may move only part of its length, say the length of 10 or 20 unit molecules, in one elementary flow process. The energy of activation for flow would then be $10\Delta E_{vap}/4$ or $20\Delta E_{vap}/4$, where $\Delta E_{vap}$ is the energy of vaporization per unit molecule. If $\Delta E_{vap}$ is say 8000 cal., these energies of activation would be 20,000 or 40,000 cal., so that the resin would soften at temperatures probably below the decomposition temperature. The energy of activation for a whole polymerized molecule of say 500 unit molecules would be 1,000,000 cal., and the resin would never soften if the whole length of the molecule had to flow in one elementary flow process.

### *Water and Associated Liquids*

Liquids composed of molecules containing OH or NH groups are those ordinarily called abnormal or associated liquids. Among the other anomalies they have much higher viscosities than would be expected from the size and structure of the molecules. This abnormally large viscosity is, of course, due to the hydrogen bond structure of these liquids. If it were not for the hydrogen bond structure $H_2O$ and $NH_3$ and $CH_4$ would all have about the same viscosity; $CH_3OH$, $CH_3NH_2$ and $CH_3CH_3$; phenol, aniline and toluene, etc. The maximum possible number of hydrogen bonds per molecule is twice the number of OH and NH groups in the molecule.

Most liquids possessing hydrogen bonds supercool to form very viscous liquids and eventually glasses, for instance, glucose, glycerol, glycol, methanol, piperine, brucine, betol, and practically all sugars and alkaloids, many phenolic compounds, lower alcohols, etc. Glucose, for instance, has five OH groups and can form as many as 10 hydrogen bonds. A large molecule like glucose would also have a large energy of vaporization due to van der Waals and dipole forces, and the total energy of activation for flow might therefore be as much as 80,000 cal.

Monohydric alcohols have a maximum coordination of two, and the fact that ethyl alcohol and water have about the same viscosity indicates that ethyl alcohol must possess a large fraction of its maximum two-coordination. Alcohols may be composed of chains




with the $R$ groups arranged roughly parallel in

the liquid, or possibly rings containing three or four molecules linked in similar fashion.

Quantitative calculation of the viscosity of these associated liquids is not possible because the relative contribution to the energy of vaporization of van der Waals, dipole, hydrogen bond and repulsive forces is not known. A study of the viscosity and its temperature variation can, however, throw some light on the question. The writer[2] was able to use viscosity data to estimate the number of hydrogen bonds per molecule in water and it was concluded that the average number of bonds was about $2\frac{1}{2}$ per molecule at 0°C, $1\frac{1}{2}$ at 50°, 1 at 100° and $\frac{1}{2}$ at 150°. Heavy water is about 27 percent more viscous than ordinary water. This is easily explained by the fact that deuterium bonds are somewhat stronger than hydrogen bonds. It would require only a very small increase in the energy of activation of flow to account for a difference in viscosity of only 27 percent. This is a good illustration of the exceptional sensitivity of the viscosity to structural differences, compared to the relative insensitivity of most physical properties.

### *Non-Newtonian Flow and the Viscosity of Rubber*

If $\sinh f\lambda_2\lambda_3\lambda/2kT$ in Eq. (2) is expanded in the most general way we have

$$\eta=\frac{\dfrac{\lambda_1 h}{\lambda^2\lambda_2\lambda_3}\dfrac{F_n}{F_a^*}e^{\Delta E_a/kT}}{1+\dfrac{1}{6}\left(\dfrac{f\lambda_2\lambda_3\lambda}{2kT}\right)^2+\dfrac{1}{120}\left(\dfrac{f\lambda_2\lambda_3\lambda}{2kT}\right)^4+\cdots}, \qquad (13)$$

which reduces to Eq. (3) when $f\lambda_2\lambda_3\lambda/2 \ll kT$. This inequality may be taken as a theoretical definition of Newtonian flow. For higher shearing stresses, however, the viscosity decreases with increasing shearing stress, which is in accord with experimental fact. This effect should occur in any flowing material, but in substances of low viscosity, turbulence sets in before the shearing stress has become high enough, and then the hydrodynamic definition of viscosity is no longer valid. Homogeneous substances of low viscosity are therefore called Newtonian liquids, because their regime of nonturbulent flow is always in the region of low shearing stress. More viscous substances will show Newtonian flow at sufficiently small stresses, but at higher stresses they will show non-Newtonian flow, i.e., a variation in the ratio rate of shear/shearing stress with shearing stress. For very viscous substances even the smallest stresses giving conveniently measurable deformations may be in the non-Newtonian range.

In cases where $f\lambda_2\lambda_3\lambda/2$ is sufficiently large the backward rate of flow can be neglected in comparison with the forward rate of flow as Eyring[1] has pointed out (see Fig. 2). For this case

$$\eta=f\lambda_1/\lambda k_f$$
$$=\frac{f\lambda_1}{\lambda}\frac{h}{kT}\frac{F_n}{F_a^*}e^{(\Delta E_a-\frac{1}{2}f\lambda_2\lambda_3\lambda)/kT}. \qquad (14)$$

This same result is obtained by making the approximation $\sinh x \cong e^x$ for $x \geqq 1.5$ in Eq. (2), as Smallwood[28] has done.

Figure 10, taken from the work of Ferry and Parks[29] on the viscosity of polyisobutylene, shows an example of flow over a wide range of shearing stress. The region of Newtonian flow appears to extend up to a torque of about 200,000 dyne-cm. The reciprocal of the slope of this "consistency" curve at any point multiplied by the constant of the instrument is the viscosity. Sometimes these viscosities which are not independent of shearing stress are called "apparent viscosities," but this development makes it clear that they are all true viscosities. However, the slope of the linear portion of the curve at small shearing

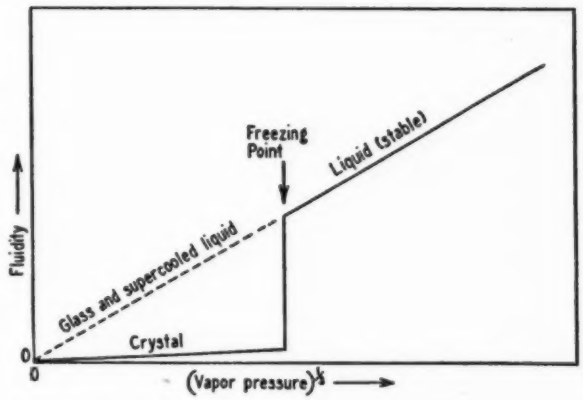



FIG. 12. Schematic representation of the fluidity and vapor pressure relations of liquids, crystals and glasses.

stresses gives the viscosity which is to be compared to the Newtonian viscosities of other liquids.

Smallwood[28] has applied Eq. (14) in the form

$$\eta = A f e^{(a-bf)/T} \qquad (15)$$

to the experimental data of Mooney[30] on the viscosity of raw rubber as a function of temperature and shearing stress. A graphical evaluation of the constants gave: $A = 1.65 \cdot 10^{-4}$, $a = 5100$, $b = 1.28 \cdot 10^{-5}$. Using these constants the equation gave good agreement with experiment for shearing stresses between 840,000 and 1,880,000 dynes/cm² and at temperatures of 70 to 140°C.

The constant $a$ is equal to $\Delta E_{vis}/R$, so that $\Delta E_{vis} = 10.1$ kcal. Raw rubber consists of linear polymers of isoprene of the order of 1000 isoprene units per chain. Such a chain would not move a whole chain length per unit flow process, but probably only the length of several isoprene units as Smallwood has suggested. The energy of vaporization of isoprene at 25° is 5.7 kcal./mole. At the higher temperatures of these experiments this energy would be less, and also the energy of vaporization per isoprene unit in rubber would be somewhat less than the energy of vaporization for an isolated isoprene molecule. If we assume then that the energy of vaporization of rubber per isoprene unit is about 4 kcal. and that $n = 4$ for this molecule, the energy of activation for flow per isoprene unit would be 1 kcal. This leads to the conclusion that during the flow process, the rubber molecule moves about 10 isoprene units.

The constant $b$ is equal to $\lambda_2\lambda_3\lambda/2k$ and substituting the experimental value we have

$$\lambda_2\lambda_3\lambda/2k = 1.28 \cdot 10^{-5},$$
$$\lambda_2\lambda_3\lambda = 3.51 \cdot 10^{-21} \text{ cc.}$$

Now assume the length of the isoprene unit, $\lambda_2$, is 5A and the thickness, $\lambda_3$, is 4A. The last paragraph showed that $\lambda = 10\lambda_2$, or $\lambda = 50$A. These values give $\lambda_2\lambda_3\lambda = 1 \cdot 10^{-21}$ cc, in fair agreement with the above experimental value. The factor of 3.5 might mean, as Smallwood suggests, that a bundle of 3 or 4 hydrocarbon chains moves as a unit.

Equation (14) represents the viscous behavior of a non-Newtonian liquid, or more accurately of a non-Newtonian state of flow. The equation has, however, both upper and lower limits of applicability. On the side of very small stresses the equation is not valid when $f\lambda_2\lambda_3\lambda/2kT < 1.5$, and on the side of large stresses the whole development is invalid when $f\lambda_2\lambda_3\lambda/2 = \Delta E_a$.

### *Flow in Crystals*

When a liquid crystallizes the viscosity usually increases by a factor of approximately $10^{12}$. The energy of activation for flow is obviously much greater in the crystal, and it seems likely that the regular structure of the crystal requires a *whole* molecule-sized hole, and therefore *all* the energy of vaporization (sublimation) to activate the flow process, compared to the $\frac{1}{3}$ or $\frac{1}{4}$ the energy of vaporization in the case of liquids. Since the best of fusion is relatively small, $\Delta E_{vis}$ for the crystal should be about 3 or 4 times $\Delta E_{vis}$ for the liquid, if the two entropies of activation for flow are the same.

The only substance for which viscosity measurements on both liquid and crystal are available is menthol, shown in Fig. 11. The slope of the liquid line gives $\Delta E_{vis} = 12{,}000$ cal. and the slope of the crystal line gives $\Delta E_{vis} = 34{,}000$ cal. Since the entropies of activation for flow will probably be different for liquid and crystal, the approximate ratio of 3 between the two values of $\Delta E_{vis}$ is satisfactory. The energies of vaporization of liquid and crystalline menthol have never been measured.

The high viscosities of glasses and crystals differ in an important respect. High viscosity in a glass is accompanied by a correspondingly small vapor pressure, while the vapor pressure of a crystal is relatively high. Fig. 12 illustrates this point. When a liquid is cooled the fluidity and the vapor pressure both decrease as discussed in an earlier section. At the freezing point the liquid may crystallize decreasing the fluidity enormously (by a factor of about $10^{12}$), while the vapor pressure remains unchanged. However, if the liquid supercools, the fluidity and the vapor pressure continue to decrease without any discontinuity (dotted line in Fig. 12), and only when the vapor pressure has become exceedingly small does the fluidity become very small.

### *Conclusion*

The concepts and relations outlined in this article can be applied to all flow phenomena, quantitatively in some cases, and only qualitatively in others. A few additional fields of application that suggest themselves are creep phenomena in metals, diffusion in both liquids and crystals, the viscosity of solutions of all sorts, including solutions of rubber, proteins, cellulose esters, and other large molecules, and the electrical conductivity of glasses, liquids and solutions containing ions.

### *Bibliography*


1. Eyring, J. Chem. Phys. **4**, 283 (1936).
2. Ewell and Eyring, J. Chem. Phys. **5**, 726 (1937).
3. Eyring, Chem. Rev. **17**, 65 (1935).
4. Eyring and Hirschfelder, J. Phys. Chem. **41**, 249 (1937).
5. Hirschfelder, Stevenson and Eyring, J. Chem. Phys. **5**, 896 (1937).
6. Hildebrand, *Solubility*, (Reinhold Publishing Co., 1936) p. 98 ff.
7. Bridgman, Proc. Am. Acad. Arts Sci. **61**, 57 (1926).
8. Bridgman, Proc. Am. Acad. Arts Sci. **66**, 185 (1931).
9. Dow, J. Wash. Acad, Sci. **24**, 516 (1934).
10. Dow, J. App. Phys. **8**, 367 (1937).
11. de Guzman, Anales soc. esp. fis. quim. **11**, 353 (1913).
12. Arrhenius, Medd. Vetenskapsakad. Nobelinst. **3**, No. 20 (1916).
13. Kendall and Monroe, J. Am. Chem. Soc. **29**, 1799 (1917).
14. Drucker, Zeits. f. physik. Chemie **92**, 287 (1917).
15. Raman, Nature **111**, 532 (1923).
16. Fulcher, J. Am. Cer. Soc. **8**, 339 (1925).
17. Dunn, Trans. Faraday Soc. **22**, 401 (1926).

17a. Busse and Karrer, quoted by Karrer, Berl and Umstätter, Zeits. f. physik. Chemie **152A**, 150 (1931).

18. Andrade, Nature **125**, 309 (1930).
19. Sheppard, Nature **125**, 489 (1930).
20. Tonomura, Bull. Chem. Soc. Japan **6**, 124 (1931).
21. Cragoe, quoted by Vinal and Craig, Nat. Bur. Stand. J. Research **10**, 781 (1933).
22. Andrade, Phil. Mag. **17**, 689 (1934).
23. Souders, J. Am. Chem. Soc. **59**, 1252 (1937).
24. Warren, Zeits. f. Krist. **86**, 349 (1933).
25. Warren and Loring, J. Am. Cer. Soc. **18**, 269 (1935).
26. Bair, J. Am. Cer. Soc. **19**, 339 (1936).
27. Taylor and Dear, J. Am. Cer. Soc. **20**, 296 (1937).
28. Smallwood, J. App. Phys. **8**, 505 (1937).
29. Ferry and Parks, Physics **6**, 356 (1935).
30. Mooney, Physics, **7**, 413 (1936).


---

*The Commissioners of the 1851 Exhibition made a good choice when they elected Rutherford as a scholar. Indeed, has any money ever been better invested? This award enabled Rutherford to go to the Cavendish Laboratory at Cambridge.*

NATURE, October 30, 1937

# Here and There

### Pennsylvania Conference

The 1938 Spring Meeting of the Pennsylvania Conference of College Physics Teachers was held April 1–2 at St. Thomas College, Scranton, Pennsylvania, by invitation of the physics departments of St. Thomas College and Marywood College.

★

### Ohio Physics Club

The Ohio Physics Club met at Kenyon College, Gambier, Ohio on Saturday, February 19, 1938. The following program was presented:

1. **Welcome to the Kenyon Campus,** by GORDON K. CHALMERS, President, Kenyon College
2. **Can Freshmen Learn Scientific Reasoning?** by C. E. HESTHAL, Ohio State University
3. **Resolution of the 4.3 mu Band of Carbon Dioxide,** by D. M. CAMERON, Ohio State University
4. **Excitation of Active Nitrogen,** by ALLEN MURRAY, Ohio State University
5. **A Simple Method of Sustaining Longitudinal Vibrations in Rods,** by E. H. JOHNSON, Kenyon College
6. **Recent Developments in the Measurement of Radio Field Strength,** by P. B. TAYLOR, Wright Field
7. **Some Impressions on the Teaching of Undergraduate Physics in England and Wales,** by G. E. OWEN, Antioch College

There was an invited paper by L. M. Heil, Research Associate of Progressive Education Association on "Evaluating the Outcomes of Physics Instruction."

★

### New Director at Kelvinator Laboratories

H. G. Chamberlin, previously connected with Kelvinator's Research Engineering Department in the air-conditioning field, has succeeded Dr. Philipp as Director of Laboratories for Kelvinator. Mr. Chamberlin, a graduate of the University of Michigan, was connected with the engineering and research departments of three of the other well-known manufacturers of household appliances before joining the Kelvinator organization.

★

### Story of Plastics

The dramatic story of the development of an idea into a material, and the development of a material into a great industry was presented at the Museum of Science and Industry, Rockefeller Center, New York, from March 1 to 31 inclusive. On view was a reproduction of a corner in the laboratory in Yonkers, New York, where, in 1907, Dr. L. H. Baekeland developed Bakelite resinoid, a discovery which marked the beginning of our modern plastics industry. The growth of the industry was presented by a series of animated and dramatic exhibits, indicating where plastics play a part in modern living. Among them were the uses of plastics in aviation, household appliances, building, business machines, abrasives, home furnishings, photography and optics, packaging, radio, machinery, fashions, automobiles, paints and varnishes, communications, amusements, music and the health of humanity.

★

### Industrial Research Laboratories

The increase in the number of research laboratories maintained by industrial concerns in the United States during the last few years has made it seem desirable to issue a new edition of the National Research Council's Bulletin, "Industrial Research Laboratories of the United States," fifth edition.

On March 25th questionnaires were mailed to the 1562 concerns which were included in the last edition of 1933, and to a large number of new concerns which are thought to maintain laboratories.

If the reader of this note is a member of a firm which maintains a laboratory where research looking towards the development and improvement of products is carried on, it is hoped that he will ascertain whether a questionnaire has been received by his company, and if not that he will request one from the Library, National Research Council, 2101 Constitution Avenue, Washington, D. C.

There is no charge for the entry in the bulletin, the only requirement being that the laboratory is undertaking research. It is desirable to have the information for the bulletin in hand as soon as possible so that the publication may appear within the current year.

★

### Necrology

Francis Charles McMath, retired engineer and bridge builder and amateur astronomer, died in Detroit, February 13th. Together with his son, Robert R. McMath, and Judge Henry S. Hulbert of Detroit, he received the John Price Wetherill Medal (1933) of the Franklin Institute for "their design and construction of novel apparatus for the making of motion pictures of astronomical substances." Among the most interesting products of the McMath Hulbert Observatory near Ann Arbor, Michigan, have been such pictures of solar prominences. Studies now under way promise to reveal much about the astrophysical processes occurring in the sun's surface.

# Innovations in Instruments

### Facsimile Printer

After ten years of laboratory research, RCA Victor engineers have announced that a simplified radio facsimile system designed to flash pictures, news bulletins and other text through the air and into the homes has been made available to radio broadcasters. The new facsimile system, which was developed by Charles J. Young, RCA Victor research engineer, and his associates, utilizes ordinary white paper, or newsprint paper, and ordinary carbon paper, at the receiving end. The width of the paper roll on which the facsimile material is printed is 8½ inches. The proposed length of each page is 12 inches on a continuous roll. Printing speed is at the rate of three feet of record per hour. A time-clock arrangement at the printer-receiver automatically turns the apparatus on and off in accordance with a predetermined transmission schedule.

The picture, drawing or text to be transmitted is placed on the roller drum of the "scanner." A beam of light travels horizontally across the page as the drum revolves. The light is reflected and focused on a sensitive photoelectric cell in the various degrees of shading corresponding to the picture. The photoelectric cell transforms the light into electrical impulses which are flashed through the air. The receiver is synchronized to the transmitter-scanner. The signals are picked up exactly as in sound broadcasting, but instead of passing through the loudspeaker, they are made to actuate the printer mechanism. Continuously feeding rolls of ordinary white paper and ordinary carbon paper are led past a metal cylinder drum, on which a single spiral of wire projects a fraction of an inch above the surface. The fluctuations in the intensity of the incoming signals press the paper and carbon together against the spiral to make marks corresponding to the light and shade of the original at the scanner.

★

### Machining World's Largest Bearing

The largest journal bearing ever made, the 317,000 pound "horseshoe" bearing which will carry the million-pound load of the world's largest telescope atop Mt. Palomar, California, is being machined at the East Pittsburgh plant of the Westinghouse Electric and Manufacturing Company. During machining the "horseshoe" is bent out of shape while its surface is machined to a glass-smooth finish so that it will be squeezed back into a perfect circle under the weight of the telescope mounting. To machine $\frac{1}{16}$" off the face of the "horseshoe" requires 63 hours. It resembles a giant steel washer 46 feet across and 53 inches thick, from which a large "U" has been cut. Because of its unusual size the bearing had to be made in three sections so that it could be shipped on flat cars. So precise are the calculations for the final maching of the bearing that it will be within five thousandths of an inch of a perfect circle when it springs into its normal position.

★

### Aerial Mapping

Aerial mapping is now reaching a stage of accuracy and speed which is destined to relieve the surveyor, with his transit and theodolite, of much of his roughest work. Jobs that would have taken him weeks can now be accomplished in a few hours by the use of a plane, a camera, and a new stereo-mapping projector which has just come from the Scientific Bureau of Bausch & Lomb Optical Co. An automatic camera in the plane, shooting at regular intervals, makes pictures a mile apart. Terrain features are thus seen from different positions in succeeding photos just as the two eyes seeing things from slightly different positions get depth perception. If the two eyes, respectively, see the views taken a mile apart, the effect is as if the mapmaker had eyes a mile apart.

To achieve this effect, the 7×9" film negatives, each covering from the usual altitude of 20,000 feet, about 13 square miles, are printed on small glass plates about the size of two special delivery stamps. The utmost exactness is required in the adjustment of the instruments since a difference of one ten-thousandth of an inch might mean a difference of feet in the field. From the glass plates the picture is projected down on a drafting table by two adjacent projectors operating in red and blue light, respectively.

With six separate adjustments on each projector set to produce exactly the same angular position that the camera occupied when making the corresponding negative, the mapper, wearing spectacles with one red and one blue lens, suddenly sees a single illusory three-dimensional model of the terrain on the table before him. The effect is so realistic that he may feel an impulse to pat the top of a smooth hill or prick his finger on a telephone pole.

To draw his map he moves across the drawing paper a fixture containing an illuminated pinhole mounted directly above a pencil. With this point of light set at a given height, the mapmaker moves the fixture about so as to keep the point in contact with the surface of the illusory ground. The line thus traced passes through all points where the ground is that high, resulting in a contour map.

The stereo-mapping projector will be used by the U. S. Army Air Corps, for military work and by government

agencies and private mapping outfits for forest surveys, estimating timber stands, laying out logging areas, locating dam sites and camps and making topographic surveys. Aerial photography combined with the new mapping instrument reveals to the engineer with uncanny accuracy the problems with which he must deal.

★

### New Photoflash Lamps

A photoflash lamp so small that eighteen can be carried in a man's coat pocket has been announced by the Westinghouse Electric and Manufacturing Company. The lamp has a filament which burns out before the light-giving element starts burning. This saves the small dry cells which photographers usually carry to set off their flashes and, more important, greatly aids in timing the flash with the opening of the shutter in high speed photography.

This vest pocket flash bulb contains metal foil and metal wire. Both burn when flashed but for a longer period than former types so that synchronization of the camera shutter with the lamp flash is easier. It is particularly a boon to users of miniature cameras having curtain shutters. The new bulb is accurately timed and is designed to have the peak of its flash timed exactly like the peak of the older types made by the same company so that cameras already synchronized to Mazda No. 20 lamps do not have to be changed.

### Speed of Deer Flies

There is a fantastic yarn about an insect that flies so fast it arrives before you hear it leave—a fly with a speed of 818 miles an hour (sound travels only 740 miles an hour) which has been disputed by Dr. Irving Langmuir, associated director of General Electric's research laboratory and Nobel prize winner.

Writing in the March 11 issue of *Science*, he has shown:

(1) The nonstreamlined—and really flat-headed—fly at such a speed would encounter a wind pressure against the head of about eight pounds per square inch, probably enough to crush the fly. $R = pd^2v^2f$ is his proof.
(2) Power consumption for maintaining a velocity of 818 miles per hour would be about one-half horsepower which, for a fly, would be a good deal.
(3) If the fly equals man in efficiency—and it doesn't seem as though the fly can be more efficient than man, who himself has a high thermodynamic efficiency—the fly must consume 1.5 times his own weight of food each second to deliver the 370 watts, or nearly half horsepower of energy, such speed requires.
(4) A piece of solder of roughly the dimensions of the fly can be seen only as a blur if, in a brightly-lighted, white-ceilinged room it is attached to a silk thread and swung so that its speed is 13 miles per hour; that at 26 miles per hour it is barely visible; that at 43 miles per hour it appears only as a faint line and that its direction of flight cannot be told; and that at 64 miles per hour the solder "fly" is invisible. (The higher speed attributed to the fly was based on a published report of an entomologist that he had seen the insects fly past him so fast they could be seen only as a blur, and that he estimated their speed as 400 yards a second, or 818 miles per hour.)
(5) Laboratory light intensity measurements and calculations also prove that an object the size of the deer fly is invisible at speeds of 64 miles per hour and above.
(6) A fly striking a person at such a speed would exert force of 310 pounds, or about four tons per square inch, and would penetrate deeply into human flesh—and the first such instance is yet to be reported; and finally
(7) Dr. Langmuir concludes a speed of 25 miles per hour is a reasonable one for the deer fly, while 800 miles per hour is utterly impossible.

# Some of the New Books

**Electrical Measurements.** FRANK A. LAWS. Second edition. Pp. 739+xiv. Figs. 476. McGraw-Hill Book Co., Inc. New York and London, 1938. Price $6.00.

The second edition of this excellent book appears in a well-proportioned volume in the series of Electrical Engineering texts. The name of the publisher assures one that this is a new book, and not merely a reprint of the 1917 edition. By using larger pages of standard size considerable new material is added without materially increasing the number of pages. Each chapter shows evidence of revision, with the omission of many illustrations of antiquated apparatus and pictures of modern instruments replacing those of older types.

The book is replete with descriptions and illustrations of electrical measuring instruments. There are chapters on the instruments for measuring current, resistance, electromotive force, power, inductance and capacitance, as well as phase meters, instrument transformers, wave form, and cable testing. New sections of ten to a dozen pages each are given to shielded resistances; the quadrant electrometer; vacuum tube voltmeters; mutual inductometers; high voltage impedance bridges; alternating current meters; transformers; and the analysis of wave form. The reader is supposed to be familiar with calculus and complex numbers, as well as the theory of alternating currents. Many references have been added, enabling one to find a more detailed discussion of each subject. As stated in the preface, the book is a compendium of information about electrical measurements and a textbook for students. It certainly is the former, but an instructor would need to make a careful selection of the pages that would be suitable for a given class of students.

Instruments for measuring electrical quantities are fully described, but the methods of procedure where the measurement is given by zero current in the detector are slighted. Thus, several types of potentiometers as made by the best makers of electrical apparatus are described and illustrated by pictures and detailed drawings of the internal electrical circuits, but the use of the potentiometer in making accurate and expeditious measurements is dismissed as "a time consuming operation, very difficult of practical realization." Power factor instruments are included, but the bridge

method of Rosa is not given. The fluxmeter is discussed, but measurements of the magnetization and hysteresis of steel are not mentioned. Many things one has seen in the text cannot be found again through the aid of the index of a dozen pages, which is inadequate for a compendium.

All in all the book is an interesting and valuable treatise, full of suggestive comments of practical application. It will be useful to anyone who is engaged in making electrical measurements.

ARTHUR W. SMITH
*University of Michigan*

---

**Low Temperature Physics.** M. and B. RUHEMAN. Pp. 313+ix. University Press, Cambridge; The Macmillan Company, New York, 1938. Price $5.00.

This book may be looked upon as an excellent volume supplementary to those "textbooks and monographs" already published. As the authors disclaim any attempt at producing what might be called a textbook one cannot perhaps complain at the parts of low temperature physics which have been almost entirely omitted from the volume. In spite of the authors' disclaimer the reading public might justifiably expect that a book with the comprehensive title of "Low Temperature Physics" should have given more space to the phenomenon of syperconductivity, the curious relations between magnetic fields and superconductivity, and the properties of the two forms of liquid helium, He I and He II.

This general criticism aside, only praise can be given for what the book does contain. Although the divisions of the book differ greatly in difficulty for the authors' "prospective readers" every question dealt with is treated with clearness and precision.

The book, consisting of fourteen chapters, is divided into four main parts: (I) that dealing with the general problems of liquefaction of gases, the equilibrium conditions at low temperatures and, briefly, the measurement of low temperatures; (II) that dealing with the solid (crystalline) state at low temperature; (III) that designated orbit and spin; and (IV) that called the free electron.

Part I gives a very satisfactory historical outline of the liquefaction of air and an adequate description of the various forms of hydrogen and helium liquefiers, ending up with Kapitza's and Simon's methods. After a short chapter on thermometry in this region the section concludes with a chapter on the theory of rectification and one on solid-liquid equilibrium in which Simon's experiments showing the vagaries of solid helium are described.

Part II is probably the most specialized of all as it deals with "x-rays and crystal structure" at low temperature, a field which the authors are peculiarly capable of treating. This x-ray work is followed by a full treatment of thermal properties of crystals at low temperatures.

In Part III the paramagnetism of certain salts is dealt with from a theoretical point of view, leading up to the experimental methods of magnetic cooling first developed by Giauque and used extensively by de Haas at Leiden and Simon at Oxford.

Part IV, a relatively short section, is given up to the thermal and electrical conductivity of metals at low temperature, a relatively small amount of attention being given to the problems of superconductivity.

E. F. BURTON
*University of Toronto*

---

**Light, Photometry and Illuminating Engineering.** WILLIAM E. BARROWS. McGraw-Hill Book Co., Inc., New York, second edition, 1938. Price $4.00.

The reviewer of a book customarily points out the good features such as size, color of binding, etc., and then exhibits his erudition by indicating a few statements in the book that are not in accordance with the conclusions of other authorities. I shall follow this formula. In Professor Barrows' book there is a great deal that can be honestly praised and there is considerable that can be condemned. Considered as a treatise on the art of modern lighting, the book is admirable; considered as a textbook of applied physics, it reminds one, unfortunately, of the Sunday supplement.

This, the second edition of *Light, Photometry and Illuminating Engineering*, constitutes a thorough revision of the first edition which was published in 1925. Approximately two-thirds of the present material is new, and the book abounds in diagrams and halftones illustrating the latest in lighting practice. One is particularly struck by the numerous illustrations of built-in lighting systems using luminous panels, lens plates, coves, louvered sources, etc., also by the information on modern light sources. Chapters VIII to XIV, inclusive, are devoted to the design of exterior and interior lighting systems, including floodlighting, lighting for sports, lighting of streets, highways and airways. Results of the recent researches on highway lighting by Reid and Chanon are included.

Turning now to the unpleasant part of my task, I find that many of the statements on both physical and physiological phases of the subject seem vague if not positively incorrect. On page 9, for instance, the Wien and Planck equations are confused; on page 23, the definition of "photon" is incorrect. The usual names of photometric concepts are used but they are used interchangeably. For example, *intensity* is expressed in candlepower on p. 38, in candlepower per square inch on p. 17, in lumens per square foot on p. 41, and in footcandles on p. 59.

To summarize, the treatise will be valuable to the practicing illuminating engineer and to the architect. From the hopelessly academic standpoint of this reviewer, however, the book would have been a better book had the first seven chapters been omitted.

PARRY MOON
*Massachusetts Institute of Technology*

---

**Flow of Homogeneous Fluids Through Porous Media.** MORRIS MUSKAT. Pp. 763; Figs. 278. McGraw-Hill Co., Inc., New York, 1937. Price $8.00.

The basis of this excellent treatise on the flow of homogeneous fluids through porous media is Darcy's law. This law simply states that the velocity of fluid is proportional to

the pressure gradient and inversely proportional to the viscosity of the fluid, the proportionality constant, the permeability, being characteristic of the porous medium. The generalization of this law together with the equation of continuity and the equation of state lead to partial differential equations which form the starting points of the practical problems treated. The problems treated are of great value to the oil and gas industry and should also be of value to the hydraulic engineer as well as to those concerned with the flow of fluids through refractory or ceramic materials.

The most extended discussion of the book deals with the problems of fluid flow in the steady state, requiring solutions of Laplace's equation. Particular stress is laid upon the illustration of the various analytical methods available for solving flow problems. This section might well be called "Exercises in Potential Theory" and should be of interest to students of potential theory. Practical engineering problems include, among others, dam seepage, irrigation and drainage problems, gravity flow systems, acid treatment of limestone wells, water coning, methods of spacing wells to control drainage across lease boundaries, and the water flooding efficiencies of various spacing patterns.

The third section of the book is concerned with the treatment of typical cases of the nonsteady state flow of liquids, requiring essentially the solution of the Fourier heat conduction equation. The problem of production decline and pressure changes in water drive fields is treated. The final section deals with the flow of gases through porous media. Typical steady state problems are solved. The analysis for nonsteady state problems is necessarily approximate because of the nonlinearity of the differential equation.

The book is clearly and effectively written. Extended nonmathematical summaries are appended to the mathematical chapters discussing the physical contents of the problems treated. Finally a complete list of the quantitative results derived in the text is given in the last appendix.

H. KRUTTER
*Pennsylvania State College*

---

**Segmental Functions Text and Tables.** C. M. SMOLEY. Pp. 428. C. K. Smoley and Sons, Scranton, Pa., 1937.

**Einfuhrung in die Quantenchemie.** H. HELLMAN. Pp. 350. Franz Deuticke, Wien and Leipzig, 1937. Price 20 gold marks.

---

### Booklets Recently Received

**The Transducer Corporation** of 30 Rockefeller Plaza, New York City, has just published a pamphlet describing "Bullet" Microphones engineered and manufactured to meet special sound requirements.

**The Applied Research Laboratories** of 1208 San Julian Street, Los Angeles, California, have issued a pamphlet describing a photoelectric densitometer for accurate quantitative analysis with the grating spectrograph.

**The General Electric Company** of Schenectady, New York has issued a four-page booklet describing their new exposure meter for movies or stills—for color or black and white.

# Among Our Authors

John A. Tennant, a New York attorney, was a friend of the late Frederic Eugene Ives for over forty years. Mr. Tennant is the editor of periodicals and yearbooks dealing with photography, in which field he is an authority. He is an Honorary Member of the Royal Photographic Society of Great Britain.

Edward Epstean, a photoengraver, received the Davanne medal of the Société Française de Photographie. He is a member of A.S.M.E. and the Graphic Arts Research Bureau and served as a delegate to the International Congress on Photography. He is a contributor to photographic journals. He has translated Eder's *Geschichte;* Fouque's, *La Verité Invention Photographie* (1867), and Potonniee's *Histoire* (1925). Recently he presented his large library of works on scientific and applied photography to Columbia University.

# Contributed Original Research

## The Reflection of Plane Wave Pulses from Plane Parallel Plates

MORRIS MUSKAT
*Gulf Research & Development Company, Pittsburgh, Pennsylvania*
(Received February 10, 1938)

A wave theory analysis is given for the reflection of wave pulses from plates, which heretofore has been treated by the methods of geometrical optics. The general theorem is proved that if the reflection coefficient for a harmonic wave system is exactly periodic in the frequency of the waves, the reflections from the plate due to an incident pulse will consist of a series of wave trains of exactly the same form as the incident pulse. This theorem may be applied to electromagnetic waves polarized in and normal to the plane of incidence, longitudinal waves in fluid media, and transverse waves polarized normal to the plane of incidence in general elastic media, when absorptive and dispersive effects can be neglected.

THE treatment generally given in texts on optics of the interference phenomena at multiple surfaces of discontinuity involves the superposition of the separate waves reflected from and refracted at the individual interfaces with amplitude and phase changes appropriate to single surfaces of discontinuity. By summing the amplitudes of these individual waves a resultant is obtained which is taken to represent the amplitudes of the reflected or transmitted waves in a steady state system in which a monochromatic plane wave of given intensity is continuously incident upon the reflecting plates.

Although this procedure gives result fully equivalent to those following from the wave theory treatment of the problem, it leaves unanswered a question which arises when the incident wave system is in the form of a pulse or finite wave train. For it is assumed in the geometrical method of treating reflection problems, on the basis of the superposition principle, that if the incident wave be in the form of a pulse it will be reflected in an undistorted manner, with intensities determined essentially by a continuous wave treatment. When, however, it is observed that the individual or monochromatic components forming the incident pulse will be reflected and refracted with amplitudes depending upon the frequency, it might be anticipated that when these components are resynthesized into the reflected and refracted pulses, they will no longer have the same *form* as the incident pulse, but will rather suffer a distortion in shape. In fact, it might be argued that if such a distortion would not take place, the uniqueness of the representation of a function by a Fourier integral would be violated. It is therefore of interest to show generally that the persistence in pulse form is after all consistent with and theoretically deducible from the wave theory treatment of the problem, and that the above mentioned paradox is only apparent.[1] This will be done in the present paper.

For these considerations no detailed specifications need be made as to the exact nature of the wave system. It may then be shown quite generally by the wave theory method of Drude or

[1] A search through the standard texts on acoustics and optics has disclosed no explicit wave treatment of the present problem of pulse propagation. Such discussions, from the point of view of seismic wave propagation, have, however, been previously published in a series of papers by K. Sezawa and K. Kanai in the Bulletins of the Earthquake Research Institute, beginning with Vol. 8 (1930), but with the restriction of the incident pulse forms to finite sinusoidal or error function wave trains. Their equivalents of the $\Phi_r$ integral of Eq. (4) below were then explicitly integrated for these special cases, but no suggestions were made as to the generality of the results or their relation to the harmonic character of $R(\omega)$.

Rayleigh that the complex amplitude of the monochromatic continuous wave of frequency $\omega$, reflected from a plane parallel plate of thickness $h$, on which is incident a unit amplitude wave at the angle $\theta_1$, may be represented by the expression:

$$R(\omega)=\frac{F'\cos p+iII'\sin p}{F\cos p+iII\sin p}, \qquad (1)^2$$

where $F'=\sigma_3-\sigma_1$; $II'=1-\sigma_1\sigma_3$;

$F=\sigma_3+\sigma_1$; $II=1+\sigma_1\sigma_3$,

$$p=\frac{h\omega\cos\theta_2}{v_2}\equiv\zeta\omega; \quad \sigma_1=\frac{a_1\tan\theta_1}{a_2\tan\theta_2}; \qquad (2)$$

$$\sigma_3=\frac{a_3\tan\theta_3}{a_2\tan\theta_2},$$

$\theta_i$ being the angles of the directions of wave propagation with the normal to the plate. The subscripts $i=1, 2, 3$, refer to the incident medium, plate, and emergence medium, and $v_2$ is the wave velocity in the plate. The constants $a$ have the following significance:

$a$ = density, for waves in fluid media; $R$ = ratio of displacements or dilatation amplitudes of reflected to incident wave.

$a$ = reciprocal of rigidity modulus for general elastic media, for transverse waves polarized normal to the plane of incidence; $R$ = ratio of displacement amplitudes of reflected to incident wave.

$a$ = magnetic permeability, for electromagnetic waves polarized in plane of incidence; $R$ = ratio of reflected to incident electric or magnetic vectors.

$a$ = dielectric constant, for electromagnetic waves polarized normal to the plane of incidence; $R$ = ratio of reflected to incident magnetic or electric vectors.

If now a wave pulse impinges on the plate, the resultant system of reflected waves may be found as follows. Supposing the incident pulse to be defined by the function:

$$\Phi=\begin{matrix} f(\delta);\ 0\le\delta\le b \\ 0;\quad \text{otherwise};\ \delta=t+s_i/v_1, \end{matrix} \qquad (3)$$

where $\Phi$ represents the amplitude of the pulse, $s_i$ the distance from the origin to the phase plane $\delta$ (cf. Fig. 1), and $v_1$ is the velocity in the incident

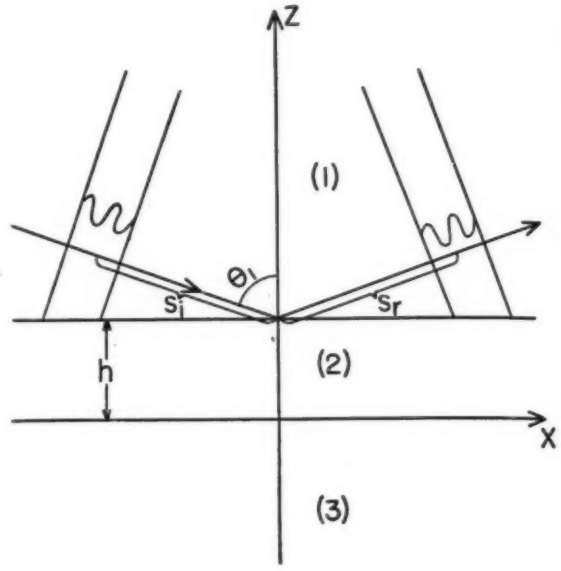



FIG. 1.

medium, the resultant wave system in the incident medium can be expressed as:

$$\Phi=\frac{1}{2\pi}\int_{-\infty}^{+\infty}d\omega\int_0^b f(g)e^{-i\omega(g-\delta)}dg + \frac{1}{2\pi}\int_{-\infty}^{+\infty}R(\omega)d\omega\int_0^b f(g)e^{-i\omega(g-\delta')}dg, \qquad (4)$$

where $\delta'=t-s_r/v_1$, and $s_r$ defines the phase planes in the reflected wave system. The latter is evidently represented by the second integral in Eq. (4).

To find the detailed nature of this second term, which we shall denote by $\Phi_r$, one may make use of either the singular or periodic properties of $R(\omega)$. Thus supposing that $R(\omega)$ has no poles in the lower half of the complex plane—which may be verified in the particular cases of practical interest without difficulty[3]—it may be shown by expressing the infinite integral for $\Phi_r$ as a closed contour over the real axis and the infinite semicircle in the lower half-plane that:

$$\Phi_r=0 \text{ for: } (s_r/v_1)>t. \qquad (5)$$

[2] A similar general representation holds for the transmitted or refracted waves. The numerators in Eq. (1) are then, of course, different, though they are still periodic in $p$. On the other hand, it is to be explicitly understood that throughout the present discussions the assumption is made that the media carrying the wave systems are neither absorptive nor dispersive.

[3] For the $R(\omega)$ of Eq. (1) this follows at once from the observation that both $F$ and $II$ are positive.

This evidently means that $\Phi_r$ represents a reflected wave system with a sharp front moving with the velocity $v_1$. Adding now the further assumption that the singularities of $R(\omega)$ in the upper half-plane are the simple poles $\omega=\omega_m$, it follows again by a contour integral evaluation of the infinite integral in $\Phi_r$ that it may be expanded as:

$$\Phi_r=i\sum_m a_m e^{i\omega_m\delta'}\int_0^{b,\,\delta'} f(g)e^{-i\omega_m g}dg+\binom{0}{1}R_\infty f(\delta'), \tag{6}$$

where the $a_m$ are the residues of $R(\omega)$ at $\omega_m$, the limit $b$, $\delta'$ means that the smaller of the two should be used, and $\binom{0}{1}$ means that 0,1 should be used as coefficients of $R_\infty$ with $b$, $\delta'$, respectively. $R_\infty$ is the constant value of $R(\omega)$ over the infinite semi-circle in the upper half-plane. Making use of the fact that the poles $\omega_m$ are distributed periodically with respect to their real components, and that the residues $a_m$ are essentially independent of $m$, it may be shown that Eq. (6) reduces to a Fourier series expansion for a series of reflected wave pulses of exactly the same form as $f(\delta)$. However, as this reduction involves making explicit use of the form of $R(\omega)$, we shall proceed directly with the following much simpler method.

Here we simply note that because of the periodic character of $R(\omega)$ it may be expanded for *all* $\omega$ as:

$$R(\omega)=\sum b_n e^{-in\zeta\omega}, \tag{7}$$

where

$$b_n=\frac{1}{2\pi}\int_0^{2\pi/\zeta} R(\omega)e^{in\zeta\omega}d\omega. \tag{8}$$

Putting Eq. (7) into the integral for $\Phi_r$ of Eq. (4) it follows at once that:

$$\begin{aligned}\Phi_r&=\frac{1}{2\pi}\sum b_n\int_{-\infty}^{+\infty}d\omega\int_0^b f(g)e^{-i\omega(g-\delta'+n\zeta)}dg\\&=\sum b_n f(\delta'-n\zeta).\end{aligned} \tag{9}$$

Equation (9) evidently represents a series of pulses of exactly the same form as the initial pulse $f(\delta)$, and spaced in such a way that $\Phi_r$ is nonvanishing only in the intervals (cf. Fig. 2):

$$0\leq t-s_r/v_1-n\zeta\leq b. \tag{10}$$

Recalling the definition of $\zeta$ in Eq. (2) it is seen that these time intervals correspond to the times of travel of the refracted wave in the interior of the plate after suffering multiple reflections within it.

Determining the $b_n$ by Eq. (8), or still more simply by a direct exponential expansion of Eq. (1), $\Phi_r$ can be finally expressed as:

$$\Phi_r=\frac{F'+H'}{F+H}f\left(t-\frac{s_r}{v_1}\right)+2\frac{H'F-F'H}{F^2-H^2}\sum_1^\infty\left(\frac{H-F}{H+F}\right)^n f\left(t-\frac{s_r}{v_1}-2n\zeta\right). \tag{11}$$

Returning now to the expressions of $F$, $F'$, $H$, $H'$ in terms of $\sigma_1$, $\sigma_3$ as given by Eq. (2), Eq. (11) takes the form:

$$\Phi_r=\frac{1-\sigma_1}{1+\sigma_1}f\left(t-\frac{s_r}{v_1}\right)-\frac{4\sigma_1}{1-\sigma_1^2}\sum_1^\infty\left(\frac{1-\sigma_1}{1+\sigma_1}\right)^n\left(\frac{1-\sigma_3}{1+\sigma_3}\right)^n f\left(t-\frac{s_r}{v_1}-2n\zeta\right). \tag{12}$$

From the form of this expression and by comparing the coefficients with the well-known Fresnel reflection coefficients, it will be seen that the first term gives the directly reflected pulse, while the $n$th term in the series represents the pulse that has suffered $2n-1$ internal reflections in the plate. It will be noted that the amplitudes of the individual pulses are exactly those that would be obtained by the conventional procedure of following the pulses as they impinge on the various interfaces and applying at each such step the appropriate reflection or refraction coefficient for a single interface. The complete equivalence of

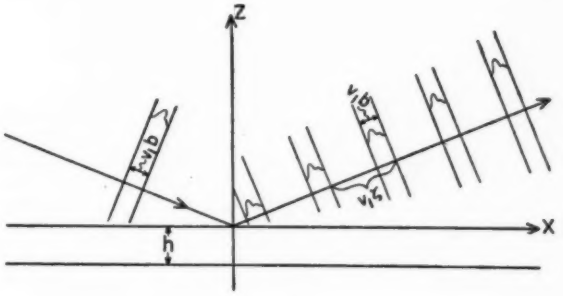



FIG. 2.

this procedure with the wave theory treatment is therefore fully established.

There yet remains the resolution of the paradox previously proposed, namely that the reflected wave system is composed of pulses of exactly the same shape as the incident pulse in spite of the fact that the reflection coefficients for the individual harmonic components depend upon the frequency. As it should be, however, this difficulty is only apparent. For from an analytical point of view the series of individually undistorted pulses gives as a whole a reflected wave system that is very definitely distorted with respect to the incident wave. Thus it will be recalled that the incident pulse is everywhere zero except when its argument $\delta$ lies in the interval between 0 and $b$, as indicated in Eq. (3). On the other hand, for the reflected wave system, which is represented by the series of Eq. (9), the nonvanishing values of $\Phi_r$ are not confined to a single interval in its fundamental variable $\delta'$, but rather they are distributed in the infinite set of intervals given by Eq. (10). Hence the resultant distribution of $\Phi_r$ over the infinite range of $\delta'$ will actually bear but little resemblance to the distribution of the incident wave function over the infinite range of its argument $\delta$. The uniqueness of the representation of a function by a Fourier integral is therefore not violated, since the second integral in Eq. (4) actually does give a different type of function than the first which represents the incident wave.

This discussion has been based upon the existence of a reflection coefficient which is strictly periodic in the frequency, as indicated in Eq. (1). Such a condition will obtain for the reflection of a wave system from a single plate for general electromagnetic waves, or for elastic waves in fluid media, or for rigid media where the wave system is composed of transverse waves polarized normal to the plane of incidence. When, however, the reflecting system is composed of two or more plates with values of $p$ (cf. Eq. (2)) which are not commensurate, or even in the case of a single plate, if the medium is a general elastic medium carrying both longitudinal and transverse waves,[4] the exact periodicity of the reflection coefficient $R(\omega)$ breaks down. Although under all such conditions it would be in principle possible to study the reflected wave system by the conventional method of tracing the pulses throughout their complete history within the various media, the multiplicity of the waves involved will clearly make such a method highly complicated. Here again, however, a wave treatment of the problem with a slight generalization of the above outlined procedure will automatically take care of the proper syntheses of the individual wave components. The detailed discussion of this problem will be published elsewhere.

The writer wishes to thank Dr. Paul D. Foote and Dr. E. A. Eckhardt of this laboratory for permission to publish this paper.

[4] In the special cases where the incident pulse is normal to the reflecting plate, Eq. (1) will still remain valid for general elastic media, provided the ratios of the tangents in $\sigma_1$ and $\sigma_3$ are set equal to unity, and $a$ is taken as the density divided by the longitudinal wave velocity for incident longitudinal waves, and the transverse wave velocity divided by the rigidity for incident transverse waves polarized in the plane of incidence. $\theta_2$ in the expression for $p$ must also, of course, be set equal to zero.

# The Effect of Pressure Upon the Elastic Parameters of Isotropic Solids, According to Murnaghan's Theory of Finite Strain *

FRANCIS BIRCH
*Harvard University, Cambridge, Massachusetts*
(Received December 20, 1937)

Murnaghan's theory of finite deformations is applied to a discussion of the effect of hydrostatic pressure upon the elastic coefficients of an isotropic body, for small strains superposed on the hydrostatic strain. Stress-strain equations for the small strains, and the equations of small motion, are shown to have the same form as those of the classical theory, with elastic parameters which depend upon the pressure. Using a form of elastic potential identical with that of the classical theory, explicit results are found for the pressure coefficients of compressibility, Young's modulus, rigidity and so on; these are compared with such experimental results as are available, with good agreement. A single-constant formula is derived which gives the volume change of such compressible materials as sodium and cesium up to the highest experimental pressure, 45,000 kg/cm², within the experimental error.

## 1. INTRODUCTION

IN spite of many attempts to extend the theory of elasticity to the general case of strains of any magnitude, a satisfactory theory has only recently been given by F. D. Murnaghan,[1] for isotropic solids. This theory rests upon a consistent use of tensor analysis and a generalized form of the principle of conservation of energy. Notwithstanding its generality, the theory may be applied to many practical problems with little more analytical difficulty than in the case of the classical theory of elasticity; the new theory embraces a whole field of effects recognized and studied experimentally, which are completely beyond the scope of the classical theory. For example, Murnaghan has shown how the observed effect of the change of compressibility with pressure is readily accounted for in terms of his theory of finite strain, with a numerical application to the case of sodium. The present paper is intended to show how this theory may be applied to several other problems, ultimately of geological interest.

The classical theory of elasticity is restricted in two ways: first, it postulates "perfect" elasticity, that is, the strains must be uniquely determined by the stresses, and inversely; second, it is in practice limited to the case of "linear" elasticity, that is, the strains must be so small that their squares and products may be neglected in the stress-strain equations. In the new theory, the first of these limitations is retained; such phenomena as "creep," hysteresis and after-working are assumed to be absent. The strains are supposed to be perfectly reversible and uniquely determined by the stresses. On the other hand, the restriction to *small* strains is removed. From these considerations, it is apparent that the new theory must find many of its most interesting and legitimate applications to problems involving high hydrostatic pressure. The "perfectly" elastic strains which may be obtained using hydrostatic pressure are far too great to permit treatment as infinitesimals; strains of the same order of magnitude produced by other stress systems are rarely reversible, and hence fall beyond the scope of any theory of "perfect" elasticity. The present discussion will be concerned chiefly with the question of *small* deformations of any type about an initial configuration corresponding to hydrostatic pressure of any magnitude; this is directly connected with recent experimental work on the effect of pressure on rigidity,[2, 3, 4] and is relevant, as a first approximation, to the problem of the velocities of propagation of elastic waves in the earth, where small strains due to the wave motion are superposed on the large, nearly hydrostatic compression under gravitational

* Paper No. 42, published under the auspices of the Committee on Geophysical Research and the Division of Geological Sciences at Harvard University.

[1] F. D. Murnaghan, Am. J. Math. **59**, 235 (1937).

[2] P. W. Bridgman, Proc. Am. Acad. Arts Sci. **63**, 401 (1929) and **64**, 39 (1929).

[3] Francis Birch, J. App. Phys. **8**, 129 (1937).

[4] Francis Birch and Dennison Bancroft, J. Geol. **46**, 59 (1938).

forces. The theory will be developed, following Murnaghan, for an isotropic body at uniform, constant temperature.

The effect of initial stress on the stress-strain relations and on wave velocities has been discussed by a number of writers without very conclusive or concordant results. Green's[5] treatment of the question was rejected by Kelvin[6] because, for one reason, it led to the conclusion that a uniform pressure, if sufficiently great, would reduce the velocity to zero. The same objection may be brought against the theory of L. Brillouin,[7] who found that the velocities in a compressed medium were given by the expressions $V_p^2=(\lambda+2\mu-P)/\rho$ and $V_s^2=(\mu-P)/\rho$, where $\lambda$ and $\mu$ are the Lamé elastic constants, $\rho$ is the density, and $P$ the pressure. Quite different expressions, leading to increased velocities at high pressures, have been given by Hencky;[8] these are based on an empirical extension of the stress-strain relations, and appear to have no general validity. Certain writers[9–13] have considered the question of initial strain in connection with various other problems dealing with the elasticity of the earth; it is likely that some of these problems will require reconsideration in terms of Murnaghan's theory.

The initial state which alone will be considered here is one corresponding to a hydrostatic pressure, of any magnitude. The equations relating small superposed strains to the additional stresses are first found for a general form of the elastic potential; the general expressions for such "measurable" coefficients as Young's modulus, the effective rigidity and the bulk modulus are then found, and shown incidentally to obey the same connecting equations as their counterparts in the classical theory of the isotropic body. The equations of small motion are formed, from which are obtained the velocities of propagation of compressional and equivoluminal waves in a compressed isotropic medium. Finally, a special form of the potential is introduced, from which numerical values for the effect of pressure on rigidity and compressibility are derived, for comparison with experimental results.

[5] G. Green, Math. Papers, p. 293.
[6] Kelvin, Baltimore Lectures, p. 230.
[7] L. Brillouin, Ann. de physique **4**, 528 (1925).
[8] H. Hencky, Phil. Mag. **14**, 254 (1932).
[9] Rayleigh, Proc. Roy. Soc. **A77**, 486 (1906).
[10] C. Chree, Phil. Mag. **32**, 233 (1891).
[11] A. E. H. Love, *The Mathematical Theory of Elasticity* (Cambridge 1927), Section 75.
[12] J. Jeans, Phil. Trans. Roy. Soc. **201**, 157 (1903).
[13] F. H. Cilley, Am. J. Sci. **11**, 269 (1901).

## 2. Equations for Small Strains Superposed on an Arbitrary Hydrostatic Compression

In the present paper, rectangular Cartesian coordinates will be used for both the initial and final states; the general formulation of Murnaghan's theory will not be required for the problems discussed here. In this special case, the difference between covariant and contravariant tensors disappears, and the notation is somewhat simplified. Murnaghan has shown that the components $T_{rs}$ of the stress tensor are given *without approximation* by the equations

$$T_{rs}=\rho(\partial\phi/\partial\epsilon_{sr}-2\epsilon_{r\beta}\partial\phi/\partial\epsilon_{\beta s}) \qquad (1)$$

where $\rho$ is the density in the strained state, $\phi$ is the potential energy or strain-energy per unit mass, and $\epsilon_{rs}$ are the components of the strain tensor. The subscripts $r$, $s$, $\beta$, represent arbitrary values of the indices $x$, $y$, $z$; repetition of a subscript indicates summation over the indices $x$, $y$, $z$; for example, $\epsilon_{x\beta}\epsilon_{\beta y}=\epsilon_{xx}\epsilon_{xy}+\epsilon_{xy}\epsilon_{yy}+\epsilon_{xz}\epsilon_{zy}$. The components $\epsilon_{rs}$ of the strain tensor are related to the displacement vector $(u, v, w)$ by the equations,

$$\epsilon_{xx}=\partial u/\partial x-\tfrac{1}{2}\{(\partial u/\partial x)^2+(\partial v/\partial x)^2+(\partial w/\partial x)^2\},$$

with two analogous equations for $\epsilon_{yy}$ and $\epsilon_{zz}$, and

$$\epsilon_{xy}=\tfrac{1}{2}\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)-\tfrac{1}{2}\left(\frac{\partial u}{\partial x}\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\frac{\partial v}{\partial y}+\frac{\partial w}{\partial x}\frac{\partial w}{\partial y}\right)$$

with two analogous equations for $\epsilon_{yz}$ and $\epsilon_{zx}$; here $(x, y, z)$ are the coordinates in the *strained* configuration of a particle initially at $(a, b, c)$. If the second order terms are neglected, these strain components reduce to those of the infinitesimal theory, except that a factor $\frac{1}{2}$ is introduced for the shearing strains by the requirement that the strain components form a tensor. The expression for the stress components differs formally from the corresponding expression in the classical theory in the terms $-2\epsilon_{r\beta}(\partial\phi/\partial\epsilon_{\beta s})$ and in the inclusion of the density $\rho$.

It is shown by Murnaghan that for an isotropic body the potential $\phi$ involves the tensor strain components only through the three scalar invariants, $I_1$, $I_2$, $I_3$, given by

$$I_1=\epsilon_{xx}+\epsilon_{yy}+\epsilon_{zz},$$

$$I_2=\epsilon_{xx}\epsilon_{yy}+\epsilon_{yy}\epsilon_{zz}+\epsilon_{zz}\epsilon_{xx}-(\epsilon_{xy}\epsilon_{yx}+\epsilon_{yz}\epsilon_{zy}+\epsilon_{xz}\epsilon_{zx}),$$

$$I_3=\epsilon_{xx}\epsilon_{yy}\epsilon_{zz}+\epsilon_{xy}\epsilon_{yz}\epsilon_{zx}+\epsilon_{yx}\epsilon_{zy}\epsilon_{xz}-(\epsilon_{xz}\epsilon_{yz}\epsilon_{zy}+\epsilon_{yy}\epsilon_{zx}\epsilon_{xz}+\epsilon_{zz}\epsilon_{xy}\epsilon_{yx}).$$

We have then for the partial differential coefficients of $\phi$,

$$\frac{\partial\phi}{\partial\epsilon_{sr}}=\frac{\partial\phi}{\partial I_1}\frac{\partial I_1}{\partial\epsilon_{sr}}+\frac{\partial\phi}{\partial I_2}\frac{\partial I_2}{\partial\epsilon_{sr}}+\frac{\partial\phi}{\partial I_3}\frac{\partial I_3}{\partial\epsilon_{sr}}.$$

Defining the quantities $\delta_{rs}$ and $\zeta_{rs}$ by the relations, $\delta_{rs}=1$ if $r=s$, 0 if $r\neq s$, and $\zeta_{ra}\epsilon_{as}=\delta_{rs}$, we have the expressions

$$\partial I_1/\partial\epsilon_{sr}=\delta_{rs},\quad \partial I_2/\partial\epsilon_{sr}=\delta_{rs}I_1-\epsilon_{rs},$$
$$\partial I_3/\partial\epsilon_{sr}=\zeta_{rs}I_3,$$

where in differentiating, we ignore the symmetry conditions, $\epsilon_{rs}=\epsilon_{sr}$. Thus the stress components are given by

$$T_{rs}=\rho\left\{(\delta_{rs}-2\epsilon_{rs})\frac{\partial\phi}{\partial I_1}+[I_1(\delta_{rs}-2\epsilon_{rs})-\epsilon_{rs}+2\epsilon_{r\beta}\epsilon_{\beta s}]\frac{\partial\phi}{\partial I_2}+I_3(\zeta_{rs}-2\delta_{rs})\frac{\partial\phi}{\partial I_3}\right\}. \quad (2)$$

For *homogeneous* strain, Murnaghan has shown that the density in the strained state $\rho$ is related to the density in the unstrained state $\rho_0$ by the *exact* expression,

$$\rho=\rho_0(1-2I_1+4I_2-8I_3)^{\frac{1}{2}}. \quad (3)$$

Now suppose that the strain components consist of two parts, one part $\epsilon$ corresponding to an arbitrary hydrostatic pressure, the other part $\epsilon_{rs}'$ to a *small* strain of any type, so that the general strain component will be $\epsilon_{rs}=\delta_{rs}\epsilon+\epsilon_{rs}'$. In the same way, the stress components will be $T_{rs}=-P\delta_{rs}+T_{rs}'$, and we wish to find the relations between $T_{rs}'$ and $\epsilon_{rs}'$.

For the invariants we find, neglecting squares and products of $\epsilon_{rs}'$,

$$\begin{aligned}I_1&=3\epsilon+I_1',\\ I_2&=3\epsilon^2+2\epsilon I_1'+I_2'\sim 3\epsilon^2+2\epsilon I_1',\\ I_3&=\epsilon^3+\epsilon^2 I_1'+\epsilon I_2'+I_3'\sim\epsilon^3+\epsilon^2 I_1',\\ I_1'&=\epsilon_{xx}'+\epsilon_{yy}'+\epsilon_{zz}',\ \text{etc.}\end{aligned} \quad (4)$$

Treating $\epsilon_{rs}'$ and hence $I_1'$ as infinitesimals, we may write

$$\frac{\partial\phi}{\partial I_i}(I)=\frac{\partial\phi}{\partial I_i}(I_0)+I_1'\frac{\partial^2\phi(I_0)}{\partial I_i\partial I_1}+2\epsilon I_1'\frac{\partial^2\phi(I_0)}{\partial I_i\partial I_2}+\epsilon^2 I_1'\frac{\partial^2\phi(I_0)}{\partial I_i\partial I_3},$$

$$\rho=\rho_0(1-2\epsilon)^{3/2}(1-(I_1'/(1-2\epsilon)),\quad i=1,2,3,$$

where $(\partial\phi/\partial I_i)(I)$ represents the value of $\partial\phi/\partial I_i$ for the general values of $I_1$, $I_2$, $I_3$ as given in (4) and $(\partial\phi/\partial I_i)(I_0)$ the value for $\epsilon_{rs}'=0$. We find also that

$$\delta_{rs}-2\epsilon_{rs}=\delta_{rs}(1-2\epsilon)-2\epsilon_{rs}',$$

$$I_1(\delta_{rs}-2\epsilon_{rs})-\epsilon_{rs}+2\epsilon_{r\beta}\epsilon_{\beta s}=\delta_{rs}(1-2\epsilon)(I_1'+2\epsilon)-\epsilon_{rs}'(1+2\epsilon),$$

$$I_3\zeta_{rs}=\delta_{rs}(\epsilon^2+\epsilon I_1')-\epsilon\epsilon_{rs}'.$$

Substituting these expressions in (2) and noting that

$$-P=\rho_0(1-2\epsilon)^{5/2}\left\{\frac{\partial\phi}{\partial I_1}(I_0)+2\epsilon\frac{\partial\phi}{\partial I_2}(I_0)+\epsilon^2\frac{\partial\phi}{\partial I_3}(I_0)\right\}. \quad (5)$$

we find six equations of the form

$$T_{rs}'=\delta_{rs}BI_1'+(A-B)\epsilon_{rs}', \quad (6)$$

two of which, in terms of $x$, $y$, $z$ are

$$\begin{aligned}T_{xx}'&=A\epsilon_{xx}'+B\epsilon_{yy}'+B\epsilon_{zz}',\\ T_{xy}'&=C\epsilon_{xy}'.\end{aligned} \quad (6')$$

The coefficients $A$, $B$ and $C$ are given by

$$A=\rho\Big\{(1-2\epsilon)\Big(\frac{\partial^2\phi}{\partial I_1^2}+4\epsilon\frac{\partial^2\phi}{\partial I_1\partial I_2}+2\epsilon^2\frac{\partial^2\phi}{\partial I_1\partial I_3}$$
$$+4\epsilon^2\frac{\partial^2\phi}{\partial I_2^2}+4\epsilon^3\frac{\partial^2\phi}{\partial I_2\partial I_3}+\epsilon^4\frac{\partial^2\phi}{\partial I_3^2}\Big)$$
$$-3\Big(\frac{\partial\phi}{\partial I_1}+2\epsilon\frac{\partial\phi}{\partial I_2}+\epsilon^2\frac{\partial\phi}{\partial I_3}\Big)\Big\},$$
$$B=\rho\Big\{(1-2\epsilon)\Big(\frac{\partial^2\phi}{\partial I_1^2}+4\epsilon\frac{\partial^2\phi}{\partial I_1\partial I_2}+2\epsilon^2\frac{\partial^2\phi}{\partial I_1\partial I_3} \qquad (7)$$
$$+4\epsilon^2\frac{\partial^2\phi}{\partial I_2^2}+4\epsilon^3\frac{\partial^2\phi}{\partial I_2\partial I_3}+\epsilon^4\frac{\partial^2\phi}{\partial I_3^2}\Big)$$
$$-\frac{\partial\phi}{\partial I_1}+(1-4\epsilon)\frac{\partial\phi}{\partial I_2}+\epsilon(1-3\epsilon)\frac{\partial\phi}{\partial I_3}\Big\},$$
$$C=\rho\Big\{-2\frac{\partial\phi}{\partial I_1}-(1+2\epsilon)\frac{\partial\phi}{\partial I_2}-\epsilon\frac{\partial\phi}{\partial I_3}\Big\},$$

where we have now put $\rho=\rho_0(1-2\epsilon)^{3/2}$; all of the partial derivatives are to be evaluated for $\epsilon_{rs}'=0$. Between these coefficients there is evidently the same relation, $A-B=C$, as in the case of small strains of an isotropic body free from initial strain; this is of course necessary, since under the special initial strain which we have used, an isotropic body must remain isotropic, and the *form* of the equations must remain unchanged. The coefficients $A$, $B$ and $C$ evidently depend upon $\epsilon$ and hence upon $P$, the initial pressure.

We may pass at once from these equations to the formulation of the various moduluses which are of interest. In the classical theory, a modulus is the ratio of a stress of a given type, to a strain of the same type, "when there is no other stress" (Love). It seems useful to extend the notion of a modulus to the present situation, where small strains are superposed upon an initial hydrostatic compression. We may also observe that the above restriction is rarely observed in practice, the initial stress in most cases being unknown. Instead of defining a modulus, in the more general case of an initial hydrostatic pressure, as the ratio of the increment of stress to the increment of strain, it seems preferable, as will appear later, to define the modulus as the ratio of the increment of stress to the *relative displacement* due to this increment of stress. The relative displacement is the ratio of two lengths, and hence is a dimensionless quantity, just as is the strain. In any experiment to determine a modulus, we measure finally a displacement; for infinitesimal strains and zero initial stress, there is no distinction between the strain and the relative displacement, so that in this limiting case, the definitions coincide.

If the total displacement from the configuration of zero stress has the components, $u=u_0+u'$, $v=v_0+v'$, $w=w_0+w'$, where $(u_0, v_0, w_0)$ is the displacement due to the hydrostatic pressure $P$ alone, we shall have the relations,

$$\begin{aligned} u_0&=-\alpha x_0, \quad x=x_0+x', \quad u'=x',\\ v_0&=-\alpha y_0, \quad y=y_0+y', \quad v'=y',\\ w_0&=-\alpha z_0, \quad z=z_0+z', \quad w'=z'. \end{aligned}$$

where $\alpha$ is a function of the pressure, but not of the coordinates. Referring back to the equations for the strains in terms of the displacements, we find $\epsilon=-\alpha-\alpha^2/2$; since $\partial u_0/\partial x=-\alpha\partial x_0/\partial x$ $=-\alpha(1-\partial u'/\partial x)$ and $\partial u/\partial x=\partial u_0/\partial x+\partial u'/\partial x$ $=-\alpha+(1+\alpha)\partial u'/\partial x$, we find for the general strain components expressions of the form $\epsilon_{xx}=\epsilon+(1+\alpha)^2\partial u'/\partial x$, whence the increments of strain are found to be of the form,

$$\begin{aligned} \epsilon_{xx}'&=(1+\alpha)^2\partial u'/\partial x,\\ \epsilon_{xy}'&=\tfrac{1}{2}(1+\alpha)^2(\partial u'/\partial y+\partial v'/\partial x). \end{aligned}$$

To find Young's modulus, we put for the stress increments

$$T_{xx}'=T', \quad T_{yy}'=T_{zz}'=0, \quad T_{xy}'=T_{yz}'=T_{zx}'=0;$$

for the strain increments, $\epsilon_{yy}'=\epsilon_{zz}'$, so that $T'=\epsilon_{xx}'(A-B)(A+2B)/(A+B)$. The small displacements may be represented by $u'=rx$, $v'=qy$, $w'=qz$. Then if Young's modulus $E$ is defined by $E=T'/(u'/x)$. This will be given by

$$E=(1+\alpha)^2(A-B)(A+2B)/(A+B).$$

In the same way, generalizing the notion of Poisson's ratio $\sigma$, we find $\sigma=B/(A+B)$. Here it makes no difference whether $\sigma$ be defined as the ratio of strain increments, or of relative displacements.

If we imagine a simple shear superposed on the compressed state, so that, for example, the equations reduce to $T_{xy}'=\epsilon_{xy}'(A-B)$,

$$\epsilon_{xy}'=\tfrac{1}{2}(1+\alpha)^2(\partial u'/\partial y+\partial v'/\partial x),$$

with $u'=\tau y$, $v'=w'=0$, then the modulus of rigidity $G$, defined as the ratio of the stress increment to the angle of shear, is given by

$$G=T_{xy}'/\tau=(1+\alpha)^2(A-B)/2.$$

From these expressions for $E$, $G$, and $\sigma$ it is readily seen that the identical relation $E/G=2(1+\sigma)$ still holds, for any pressure. In the infinitesimal theory, there are two other identities, $E=3K(1-2\sigma)$ and $G/3K=(1-2\sigma)/(2+2\sigma)$, where $K$ is the bulk modulus, the reciprocal of the compressibility. It is necessary here to define carefully the compressibility which we intend to use. There are two definitions which are in common use. The "ordinary" compressibility $\beta^*$ is given by $-(1/V_0)(\partial V/\partial P)$, where $V_0$ is the volume at zero pressure. The other, which we may call the "true" compressibility $\beta$, is $-(1/V)(\partial V/\partial P)$; here $V$ is the volume at the pressure at which the derivative $\partial V/\partial P$ is taken. Both of these are functions of the pressure, and for most solids differ by only a few percent up to pressures of the order of $10^4$ kg/cm$^2$. It is the true compressibility which is required in the above relations. Its reciprocal, the true bulk modulus $K$, may be found from the general expression for small strain about an arbitrary hydrostatic compression by setting $T_{rs}'=-\delta_{rs}P'$, and $\epsilon_{rs}'=\delta_{rs}\epsilon'$, where $P'$ is a small increment of pressure. Since $u_0=-\alpha x_0=x_0-a$, we have $x_0(1+\alpha)=a$ ($a$ being the initial $x$-coordinate of the particle which moves to $x_0$ for a pressure $P$), giving, for small changes of pressure,

$$\frac{\partial x_0}{\partial P}=-\frac{x_0}{1+\alpha}\frac{\partial \alpha}{\partial P}.$$

Then since we may put $V=x_0^3$, the material being isotropic, we have

$$\frac{1}{V}\frac{\partial V}{\partial P}=\frac{3}{x_0}\frac{\partial x_0}{\partial P}=-\frac{3}{1+\alpha}\frac{\partial \alpha}{\partial P}.$$

The total strain $\epsilon=-\alpha-\alpha^2/2$, so for small variations of pressure,

$$\frac{\partial \epsilon}{\partial P}=-(1+\alpha)\frac{\partial \alpha}{\partial P},$$

and

$$\frac{1}{V}\frac{\partial V}{\partial P}=-\frac{1}{K}=\frac{3}{(1+\alpha)^2}\frac{\partial \epsilon}{\partial P}=-\beta.$$

The same result may be obtained more directly by differentiation of Murnaghan's general expression for the ratio of volumes in the case of hydrostatic pressure; this is given by $V_0/V=(1-2\epsilon)^{3/2}$, whence

$$\frac{1}{V}\frac{\partial V}{\partial P}=\frac{3}{1-2\epsilon}\frac{\partial \epsilon}{\partial P}=\frac{3}{(1+\alpha)^2}\frac{\partial \epsilon}{\partial P}.$$

Identifying $\partial\epsilon/\partial P$ with the value of $\epsilon'/P'$ found from the stress-strain equations, we have for the true bulk modulus $K$ the expression

$$\frac{1}{K}=-\frac{1}{V}\frac{\partial V}{\partial P}=\frac{3}{(1+\alpha)^2(A+2B)}.$$

and the connecting equations are seen to retain the same form as in the classical theory, with

$$\begin{aligned} E&=(1+\alpha)^2(A-B)(A+2B)/(A+B),\\ G&=(1+\alpha)^2(A-B)/2,\\ K&=(1+\alpha)^2(A+2B)/3,\\ \sigma&=B/(A+B). \end{aligned} \tag{8}$$

The equations of motion may be written in the general form,

$$\frac{\partial T_{xx}}{\partial x}+\frac{\partial T_{xy}}{\partial y}+\frac{\partial T_{xy}}{\partial z}+\rho X=\rho f_x, \text{ etc.} \tag{9}$$

where $(X, Y, Z)$ and $(f_x, f_y, f_z)$ are the components of the body force and of the acceleration, respectively. If we write $T_{rs}=-\delta_{rs}P+T_{rs}'$ and use the relations (6), then these equations take the form

$$\frac{\partial}{\partial x}[-P+A\epsilon_{xx}'+B\epsilon_{yy}'+B\epsilon_{zz}']+\frac{\partial}{\partial y}[(A-B)\epsilon_{xy}']+\frac{\partial}{\partial z}[(A-B)\epsilon_{xz}']+\rho X=\rho\frac{\partial^2 u'}{\partial t^2},$$

with

$$\epsilon_{xx}'=(1+\alpha)^2\frac{\partial u'}{\partial x},\quad \epsilon_{xy}'=\tfrac{1}{2}(1+\alpha)^2\left(\frac{\partial u'}{\partial y}+\frac{\partial v'}{\partial x}\right), \text{etc.},$$

for *small* displacements about the position of equilibrium corresponding to the pressure $P$. In the case of a homogeneous strain, where the body forces and pressure gradients are zero, we

have the simpler expressions,

$$\frac{A-B}{2}\nabla^2(u', v', w')+\frac{A+B}{2}\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$$

$$\times\left(\frac{\partial u'}{\partial x}+\frac{\partial v'}{\partial y}+\frac{\partial w'}{\partial z}\right)=\frac{\rho}{(1+\alpha)^2}\frac{\partial^2}{\partial t^2}(u', v', w'). \quad (10)$$

Manipulating these equations in the usual way, we obtain the equations,

$$\left.\begin{aligned} A(1+\alpha)^2\nabla^2\theta&=\rho\frac{\partial^2\theta}{\partial t^2} \\ \text{and} \\ \frac{A-B}{2}(1+\alpha)^2\nabla^2(\bar{\omega}_x', \bar{\omega}_y', \bar{\omega}_z')&\\ =\rho\frac{\partial^2}{\partial t^2}(\bar{\omega}_x', \bar{\omega}_y', \bar{\omega}_z')&, \end{aligned}\right\} \quad (11)$$

where

$$\theta=\frac{\partial u'}{\partial x}+\frac{\partial v'}{\partial y}+\frac{\partial w'}{\partial z} \quad \text{and} \quad \bar{\omega}_x'=\left(\frac{\partial w'}{\partial y}-\frac{\partial v'}{\partial z}\right), \quad \text{etc.}$$

Thus the velocity of waves of dilation or compression, without rotation, will be

$$V_P=[A(1+\alpha)^2/\rho]^{\frac{1}{2}}=\left[\frac{E}{\rho}\frac{1-\sigma}{(1+\sigma)(1-2\sigma)}\right]^{\frac{1}{2}},$$

and the velocity of waves of rotation, without volume change, $V_S=[(A-B)(1+\alpha)^2/2\rho]^{\frac{1}{2}}=(G/\rho)^{\frac{1}{2}}$. These relations hold for an initially isotropic body under hydrostatic pressure, and at uniform temperature. The velocities are given by equations of the same form as in the classical theory, but the coefficients depend upon the pressure. If as in the earth, we may regard the pressure as entirely due to the body forces, then the pressure gradients and body forces mutually cancel out; but since the pressure and density are now functions of the coordinates, the equations (10) will be only approximate, valid for small distances and for waves sufficiently short. In general, the medium will now be neither homogeneous nor isotropic; waves initially plane and harmonic will be distorted and give rise to waves of higher frequencies; the velocities will depend upon the amplitudes and the frequencies. The discussion of this problem will not be attempted in the present paper.

## 3. Special Form of Potential Energy Function

Little more progress can be made without adopting a definite form for the function $\phi$, the potential energy per unit mass. This has been shown to be a function of the strain invariants alone, $I_1$, $I_2$, $I_3$, for the case of an isotropic body at constant, uniform temperature. The function $\phi$ may thus be expanded, as shown by Murnaghan, as a power series in the strain components, in the form

$$\phi=aI_1+bI_1^2+cI_2+dI_1^3+eI_1I_2+fI_3+\cdots.$$

If the stress is zero in the unstrained state, for which $\epsilon_{rs}=0$, then $a=0$, and adopting a notation to agree with Murnaghan and the classical theory, we write,

$$\rho_0\phi=((\lambda+2\mu)/2)I_1^2-2\mu I_2 +lI_1^3+mI_1I_2+nI_3+. \quad (12)$$

Here $\lambda$ and $\mu$ are the usual Lamé elastic coefficients; for small strains, the equations will reduce to the classical ones. If we keep only the second-order terms in the expansion of $\rho_0\phi$ we shall have

$$\rho_0\phi=((\lambda+2\mu)/2)I_1^2-2\mu I_2.$$

In this case,

$$\rho_0(\partial\phi/\partial I_1)=(\lambda+2\mu)I_1, \quad \rho_0(\partial\phi/\partial I_2)=-2\mu,$$

$$\rho_0(\partial^2\phi/\partial I_1^2)=(\lambda+2\mu)$$

and the coefficients $A$, $B$ and $C$ of Eqs. (6) become

$$A=(1-2\epsilon)[\lambda+2\mu-\epsilon(11\lambda+10\mu)],$$

$$B=(1-2\epsilon)[\lambda-\epsilon(5\lambda+2\mu)]$$

$$C=(1-2\epsilon)[2\mu-\epsilon(6\lambda+8\mu)]=A-B.$$

The other coefficients are now given by

$$E=(1-2\epsilon)^{5/2}\frac{[\mu-\epsilon(3\lambda+4\mu)]}{\lambda+\mu-\epsilon(8\lambda+6\mu)}(3\lambda+2\mu)(1-7\epsilon),$$

$$G=(1-2\epsilon)^{5/2}[\mu-\epsilon(3\lambda+4\mu)],$$

$$K=(1-2\epsilon)^{5/2}(1-7\epsilon)\frac{3\lambda+2\mu}{3},$$

$$\sigma=\frac{1}{2}\frac{\lambda-\epsilon(5\lambda+2\mu)}{\lambda+\mu-\epsilon(8\lambda+6\mu)},$$

$$V_P=\left\{\frac{(1-2\epsilon)^{5/2}}{\rho}[\lambda+2\mu-\epsilon(11\lambda+10\mu)]\right\}^{\frac{1}{2}},$$

$$V_S=\left\{\frac{(1-2\epsilon)^{5/2}}{\rho}[\mu-\epsilon(3\lambda+4\mu)]\right\}^{\frac{1}{2}}.$$

The relation with the pressure may be found from the additional equation,

$$-P=(1-2\epsilon)^{5/2}\epsilon(3\lambda+2\mu).$$

For comparison with most laboratory measurements, it is sufficient to retain only the first powers of $\epsilon$, or of $P$. To this approximation, the various coefficients are as follows, where the subscript zero refers to the state of zero pressure:

$$\left.\begin{aligned}
&A=\lambda+2\mu-\epsilon(14\lambda+16\mu),\\
&B=\lambda-\epsilon(8\lambda+2\mu),\\
&C=2\mu-2\epsilon(3\lambda+7\mu),\\
&E=E_0\left\{1+\frac{P}{\mu}\frac{(3\lambda+5\mu)(\lambda+2\mu)}{(3\lambda+2\mu)(\lambda+\mu)}\right\},\\
&G=G_0\left\{1+P\beta_0\left(\frac{\lambda}{\mu}+3\right)\right\},\ \beta_0=\frac{1}{K_0}=\frac{3}{3\lambda+2\mu},\\
&K=K_0(1-12\epsilon),\\
&\sigma=\sigma_0\left\{1-\frac{P}{\lambda}\frac{\lambda-\mu}{\lambda+\mu}\right\},\\
&V_P=V_P^0\left\{1+\frac{P}{3\lambda+2\mu}\frac{13\lambda+14\mu}{\lambda+2\mu}\right\}^{\frac{1}{2}},\\
&V_S=V_S^0\left\{1+\frac{P}{\mu}\frac{3\lambda+6\mu}{3\lambda+2\mu}\right\}^{\frac{1}{2}},\\
&P=-\epsilon(3\lambda+2\mu).
\end{aligned}\right\}(13)$$

Thus Murnaghan's theory affords pressure coefficients of rigidity, compressibility, Young's modulus and so on, even when the development of the potential is restricted to the terms involving the two classical parameters, $\lambda$ and $\mu$ alone.

The validity of the approximate equations (13) depends upon the magnitude of $\epsilon$; this is approximately $-P/(3\lambda+2\mu)$. The parameters $\lambda$ and $\mu$ are of the order of $10^5$ to $10^6$ kg/cm² for most solids; the pressures in general laboratory use are of the order of $10^4$ kg/cm². Thus $\epsilon$ is usually not greater than $10^{-2}$, and the equations (13) should be adequate in many cases to represent the observed effects.

## 4. Comparison with Experiment

There are, as yet, no experiments which give directly Young's modulus or the velocity of longitudinal waves for a compressed solid. The change of rigidity with pressure has been measured by Bridgman,[2] and the change of frequency of torsional vibrations of rods with pressure, which also gives the change of rigidity, by Birch,[3] and by Birch and Bancroft.[4] In Table I are given the experimental values for the coefficient $(1/G_0)\cdot(\partial G/\partial P)$, at 30°C, and the values calculated from the approximate Eq. (13), $(1/G_0)(\partial G/\partial P)=\beta_0(\lambda/\mu+3)$. The values given for the rocks are all for a pressure of 4000 kg/cm²; initial or low pressure values for aggregates of this kind depend largely upon structural looseness, whereas the present theory is concerned only with compact material.

The reasons for restricting this theory to compact aggregates may be briefly summarized. We have used the result, true for compact aggregates of individually isotropic components or of cubic crystals, that a hydrostatic pressure exerted over the boundary of a specimen produces a homogeneous, single-valued strain throughout the interior. But for "loose" aggregates, in which the component parts do not fit together perfectly, in which there are empty spaces or "pores," hydrostatic pressure over the boundary of a specimen may produce stresses in the individual parts which are far from hydrostatic, and the assumption of homogeneous or even continuous strain may be far from justified. In such aggre-

TABLE I. *Calculated and observed pressure coefficients of rigidity.*

| | $\sigma$ | $\lambda/\mu+3$ | $\beta\cdot10^{7}$ CM²/KG | $(1/G_0)(\partial G/\partial P)\cdot10^{6}$ | | REFERENCE |
|---|---|---|---|---|---|---|
| | | | | CALC. | OBS. | |
| Steel | 0.288 | 4.35 | 5.87 | 2.56 | 2.2 | 2 |
| | | | | | 2.36 | 3 |
| Nickel | .336 | 5.05 | 5.29 | 2.67 | 1.8 | 2 |
| | | | | | 1.82 | 3 |
| Copper | .364 | 5.66 | 7.19 | 4.07 | 2.71 | 3 |
| Aluminum | .355 | 5.44 | 13.43 | 7.32 | 7.61 | 3 |
| Granite, Quincy | .229 | 3.85 | 18.8 | 7.25 | 6 to 11 | 4 |
| Diabase, Maryland | .281 | 4.28 | 11.4 | 4.88 | 4 to 9 | 4 |
| Hypersthenite | .230 | 3.85 | 9.4 | 3.62 | 4 | 4 |
| Dunite | .262 | 4.10 | 8.2 | 3.3 | 4 to 9 | 4 |

gates, the stresses will be concentrated upon the areas of contact between the parts, and zero over the surfaces not in contact. The changes of elastic parameters of such aggregates at low pressures are due mainly to changes of the areas of contact, and may be of an entirely different order of magnitude from those produced by the same pressure in a compact aggregate of the same material. On the other hand, as the results of Table I indicate, such aggregates may be made sufficiently compact by the application of high pressure to give reasonably good agreement with the predictions of the theory for isotropic homogeneous material.

The materials for which values are given in Table I are all merely quasi-isotropic; that is, they are all aggregates of crystals which are not individually isotropic. If the crystals are numerous enough in the sample studied, and are oriented at random, the aggregate will show no directional properties and may be treated as macroscopically isotropic. It is clear that in all cases, the calculated and measured coefficients are of the same order of magnitude, with surprisingly good numerical agreement.

For a number of glasses, negative pressure coefficients of rigidity and of compressibility have been found.[2, 3] This is unaccountable so long as only two parameters are used in the development of the potential. If additional elastic coefficients are introduced, so that the third powers of the strain components are retained, in the way indicated by Murnaghan, it becomes possible to have negative pressure coefficients of rigidity, and also compressibilities which increase rather than decrease with pressure. It is not possible, however, to evaluate the new coefficients with the present data.

In connection with the effect of pressure on compressibility, it is interesting that Murnaghan's theory leads to a result which resembles an empirical relationship suggested by L. H. Adams.[14] If only two parameters, $\lambda$ and $\mu$ are introduced, the change of volume with pressure may be found exactly from the two expressions, $V_0/V=(1-2\epsilon)^{\frac{1}{2}}$, $-P=(3\lambda+2\mu)\epsilon(1-2\epsilon)^{5/2}$. If we consider the "ordinary" compressibility $\beta^*$ defined by $-(1/V_0)(\partial V/\partial P)=\beta^*$, then this is $3/[(3\lambda+2\mu)(1-2\epsilon)^4(1-7\epsilon)]$ and the coefficient

$$\partial\beta^*/\partial P=-3(15-70\epsilon)/[(3\lambda+2\mu)^2 \times(1-2\epsilon)^{13/2}(1-7\epsilon)^2].$$

For small values of $\epsilon$ we have the approximate relation $\partial\beta^*/\partial P=-5\beta^{*2}$. Now the measured changes of volume are given by Bridgman and other experimenters in the form $-\Delta V/V_0=aP-bP^2$; thus $\beta^*=a-2bP$, and $\partial\beta^*/\partial P=-2b$, so we should have the general relation, $2b=5a^2$. In Table II are listed the experimental values

TABLE II. *Change of compressibility with pressure.* $-\Delta V/V_0=aP-bP^2$, *P in kg/cm².*

| | BRIDGMAN, 30°C 0 TO 12,000 KG/CM² | | EBERT, 20°C 0 TO 5000 KG/CM² | |
|---|---|---|---|---|
| | $5a^2\cdot10^{12}$ | $2b\cdot10^{12}$ | $5a^2\cdot10^{12}$ | $2b\cdot10^{12}$ |
| Fe | 1.7 | 4.7 | 1.7 | 2.6* |
| Cu | 2.6 | 6.0 | 2.6 | 5.1 |
| Ag | 4.9 | 10.1 | 5.1 | 3.6* |
| Al | 9.1 | 12.4 | 9.1 | 6.0* |
| Pb | 28.1 | 40.6 | 28.8 | 27.8* |

* Values for single crystals. All others are for polycrystalline rods.

[14] L. H. Adams, J. Wash. Acad. Sci. **17** 29 (1927).

of $5a^2$ and of $2b$ for a few metals, according to Bridgman,[15] and Ebert.[16] The measured change of compressibility is in general larger than the theoretical one, for the polycrystalline materials, and less than the theoretical one, except for iron, in the case of the single crystals. This difference between the results for aggregates and for single crystals suggests that even polycrystalline metals may be somewhat "loose" aggregates, though of course to a much smaller degree than rocks. The applicability of this theory to single crystals is open to question, but all of the crystals of Table II are cubic, and consequently isotropic so far as deformation under hydrostatic pressure is concerned. In all cases, the correct order of magnitude of the change of compressibility with pressure is predicted; but the resources of this theory are by no means exhausted, since we have restricted ourselves to two parameters; if a more general expansion for $\phi$ is used, closer agreement with experiment can of course be obtained by adjusting the additional constants. The simple relation, $2b=5a^2$, will then no longer be expected to hold. The curve which Adams drew through a number of experimental points, showing $\partial\beta^*/\partial P$ as function of $\beta^*$, is fairly closely represented by $b=5a^2$, for $\beta^*$ less than about $30\cdot10^{-7}$, rather than by $2b=5a^2$. This relation is also suggested, for the materials with low compressibilities, by the values of Table II; these are, however, the materials for which $b$ is known with the least precision.

The remarkable power of this theory is also demonstrated in its treatment of the results of measurements at the very high pressures which have recently been utilized by Bridgman. Thus Murnaghan has shown how the change of volume of sodium to 20,000 kg/cm² may be represented, with an average deviation of about one-tenth of one percent, by a two-constant formula; in this pressure range, the density of sodium is increased by nearly 20 percent. It is worth while reconsidering this case as well as adding another even more impressive example, for which the experimental data have become available since the appearance of Murnaghan's paper.

The general expression for the pressure is given by (5) as

$$-P=\rho_0(1-2\epsilon)^{5/2}\{\partial\phi/\partial I_1+2\epsilon\partial\phi/\partial I_2+\epsilon^2\partial\phi/\partial I_3\}.$$

Now if the expansion (12) is adopted for $\rho_0\phi$ we have, including the terms of the third degree in $\epsilon$, and writing $f=-\epsilon$,

$$+P=f(1+2f)^{5/2}\{3\lambda+2\mu f(27l+9m+n)\}. \quad (14)$$

If we take the approximate expression, $1+5f$ for $(1+2f)^{5/2}$, this becomes Murnaghan's formula, $P=gf+hf^2$, with $g=3\lambda+2\mu$, $h=15\lambda+10\mu-27l-9m-n$. Since however the expression $\rho=\rho_0(1+2f)^{3/2}$, from which arises the term $(1+2f)^{5/2}$ in (5), is *exact*, and does not depend upon the expansion adopted for $\rho_0\phi$, it seems preferable to retain the complete term $(1+2f)^{5/2}$. If this is done, it is possible to account for the volume changes, not only for sodium to 20,000 kg/cm², but also for sodium and cesium in the unprecedented pressure range, 0–45,000 kg/cm², in terms of *single-constant* formulas. If the coefficients of the third degree terms, $l$, $m$ and $n$ are negligibly small, then (14) becomes simply,

$$P=(3\lambda+2\mu)f(1+2f)^{5/2}, \quad =gf(1+2f)^{5/2}. \quad (15)$$

Thus we have a single arbitrary constant $g$ available for fitting the experimental data. In Table III, this expression is used for calculating the pressures required to give the observed volume changes for sodium[17] to 20,000 kg/cm²

TABLE III. *Compression of sodium.*

| $\Delta V/V_0$ | $f$ | $P$ IN KG/CM² | |
|---|---|---|---|
| | | MEASURED | CALCULATED |
| 0.0295 | 0.0101 | 2000 | 2002 |
| .0552 | .0193 | 4000 | 4002 |
| .0779 | .0278 | 6000 | 6002 |
| .0981 | .0356 | 8000 | 7965 |
| .1165 | .0430 | 10000 | 9968 |
| .1332 | .0500 | 12000 | 11968 |
| .1488 | .0567 | 14000 | 13989 |
| .1632 | .0631 | 16000 | 16019 |
| .1767 | .0692 | 18000 | 18046 |
| .1894 | .0751 | 20000 | 20098 |

[15] P. W. Bridgman, Proc. Am. Acad. Arts Sci. **70**, 285 (1935).
[16] H. Ebert, Physik. Zeits. **36**, 388 (1935).
[17] P. W. Bridgman, Proc. Nat. Acad. Sci. **21**, 109 (1935).

and in Table IV, the same is done for the new measurements on cesium.[18]

TABLE IV. *Compression of cesium.*

| $-\Delta V/V_0$ | $V/V_0$ | $f$ | PRESSURE, IN KG/CM² | |
|---|---|---|---|---|
| | | | MEASURED | CALCULATED |
| 0.182 | 0.818 | 0.0716 | 5,000 | 4,980 |
| .271 | .729 | .1172 | 10,000 | 9,876 |
| .326 | .674 | .1504 | 15,000 | 14,450 |
| .372 | .628 | .1818 | 20,000 | 19,650 |
| .398 | .602 | .2013 | 23,000 | 23,340 |
| .414 | .586 | .2135 | 25,000 | 25,850 |
| .432 | .568 | .2290 | 30,000 | 29,260 |
| .458 | .542 | .2522 | 35,000 | 34,850 |
| .481 | .519 | .2742 | 40,000 | 40,710 |
| .501 | .499 | .2947 | 45,000 | 46,710 |

The average deviation of the calculated from the measured pressures in the case of sodium is roughly two-tenths of one percent, thus slightly greater than for the two-constant formula used by Murnaghan, but still well within the probable experimental error. The constant $g$ has the value, $1.886 \cdot 10^5$ kg/cm², which gives for the initial compressibility, $\beta_0 = 3/(3\lambda + 2\mu) = 3/g = 1.59 \cdot 10^{-5}$ cm²/kg.

In the case of cesium, the observed data show a polymorphic transition, accompanied by a very small volume change, at 23,000 kg/cm²; this has been smoothed out by subtracting the relative volume change at the transition, (0.006) from the values of $-\Delta V/V_0$ for the pressures above the transition pressure, but some irregularity is doubtless introduced by the existence of this effect. In Table IV are given the smoothed volume changes, the measured pressures and the pressures calculated by the formula, $P = gf(1+2f)^{5/2}$, with $g = 49{,}780$ kg/cm². The initial compressibility is thus $\beta_0 = 3/g = 6.03 \cdot 10^{-6}$ cm²/kg. The average deviation of the calculated from the measured pressures is two percent, again no greater than the probable error for this pressure range. The maximum deviation is less than four percent, at the pressure of 45,000 kg/cm².

Bridgman[18] has also measured the compression of sodium to 45,000 kg/cm². The new values, given only for 5,000 kg/cm² intervals, are somewhat different from those of Table III, by as much as 3 percent at 20,000. Thus, using the value of $g$, $1.886 \cdot 10^5$, which gave agreement with the measurements in the lower pressure range, we find with the new data pressures as much as 8 percent too low at 45,000. By choosing a new constant $g$ for the new measurements, the average deviation of calculated from measured pressures may be reduced to less than 3 percent, which is comparable with the experimental error.

Consequently, it is unnecessary to take into account the terms of the third order in the expansion of $\rho_0\phi$ even when, at least in the case of cesium, the density has been *doubled*. The terms of the second order alone, involving only the familiar elastic constants $\lambda$ and $\mu$, are adequate, when treated in accordance with Murnaghan's theory, to give a satisfactory account of the volume changes of the most compressible solids, over the greatest pressure range now available.

In a second paper, it is intended to apply some results of Murnaghan's theory to several questions concerning the interior of the earth. The writer is very greatly indebted to Professor P. W. Bridgman for discovering several serious errors in the original manuscript of the present paper, and for suggesting a much improved method of presentation.

[18] P. W. Bridgman, Proc. Am. Acad. Arts Sci. **72**, 207 (1938).